PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

PF-0459 US
09/002,485
Reference No. 6

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C07K A2

(11) International Publication Number: WO 99/46281

(43) International Publication Date: 16 September 1999 (16.09.99)

(21) International Application Number: PCT/US99/05028

(22) International Filing Date: 8 March 1999 (08.03.99)

(30) Priority Data:

60/077,450	10 March 1998 (10.03.98)	US
60/077,632	11 March 1998 (11.03.98)	US
60/077,641	11 March 1998 (11.03.98)	US

(Continued on the following page)

(71) Applicant *(for all designated States except US)*: GENENTECH, INC. [US/US]; One DNA Way, South San Francisco, CA 94080 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: WOOD, William, I. [US/US]; 35 South Down Court, Hillsborough, CA 94010 (US). GODDARD, Audrey [CA/US]; 110 Congo Street, San Francisco, CA 94131 (US). GURNEY, Austin [US/US]; 1 Debbie Lane, Belmont, CA 94002 (US). YUAN, Jean [CN/US]; 176 West 37th Avenue, San Mateo, CA 94403 (US). BAKER, Kevin, P. [GB/US]; 1115 South Grant Street, San Mateo, CA 94402 (US). CHEN, Jian [CN/US]; 1860 Ogden Drive #4, Burlingame, CA 94010 (US).

(74) Agents: KRESNAK, Mark, T. et al.; Genentech Inc., 1 DNA Way, South San Francisco CA 94080–4990 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

Without international search report and to be republished upon receipt of that report.

(54) Title: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME

(57) Abstract

The present invention is directed to novel polypeptides and to nucleic acid molecules encoding those polypeptides. Also provided herein are vectors and host cells comprising those nucleic acid sequences, chimeric polypeptide molecules comprising the polypeptides of the present invention fused to heterologous polypeptide sequences, antibodies which bind to the polypeptides of the present invention and to methods for producing the polypeptides of the present invention.

(30) 60/077,450	10 Mar/mar 1998 (10.03.1998)	US	(30) 60/081,049	8 Apr/avr 1998 (08.04.1998)	US	(30) 60/084,414	6 May/mai 1998 (06.05.1998)	US
(30) 60/077,632	11 Mar/mar 1998 (11.03.1998)	US	(30) 60/081,195	9 Apr/avr 1998 (09.04.1998)	US	(30) 60/084,640	7 May/mai 1998 (07.05.1998)	US
(30) 60/077,641	11 Mar/mar 1998 (11.03.1998)	US	(30) 60/081,203	9 Apr/avr 1998 (09.04.1998)	US	(30) 60/084,639	7 May/mai 1998 (07.05.1998)	US
(30) 60/077,649	11 Mar/mar 1998 (11.03.1998)	US	(30) 60/081,229	9 Apr/avr 1998 (09.04.1998)	US	(30) 60/084,637	7 May/mai 1998 (07.05.1998)	US
(30) 60/077,791	12 Mar/mar 1998 (12.03.1998)	US	(30) 60/081,838	15 Apr/avr 1998 (15.04.1998)	US	(30) 60/084,643	7 May/mai 1998 (07.05.1998)	US
(30) 60/078,004	13 Mar/mar 1998 (13.03.1998)	US	(30) 60/081,955	15 Apr/avr 1998 (15.04.1998)	US	(30) 60/084,598	7 May/mai 1998 (07.05.1998)	US
(30) 09/040,220	17 Mar/mar 1998 (17.03.1998)	US	(30) 60/081,952	15 Apr/avr 1998 (15.04.1998)	US	(30) 60/084,600	7 May/mai 1998 (07.05.1998)	US
(30) 60/078,886	20 Mar/mar 1998 (20.03.1998)	US	(30) 60/081,817	15 Apr/avr 1998 (15.04.1998)	US	(30) 60/084,627	7 May/mai 1998 (07.05.1998)	US
(30) 60/078,910	20 Mar/mar 1998 (20.03.1998)	US	(30) 60/082,569	21 Apr/avr 1998 (21.04.1998)	US	(30) 60/085,339	13 May/mai 1998 (13.05.1998)	US
(30) 60/078,939	20 Mar/mar 1998 (20.03.1998)	US	(30) 60/082,568	21 Apr/avr 1998 (21.04.1998)	US	(30) 60/085,338	13 May/mai 1998 (13.05.1998)	US
(30) 60/078,936	20 Mar/mar 1998 (20.03.1998)	US	(30) 60/082,700	22 Apr/avr 1998 (22.04.1998)	US	(30) 60/085,323	13 May/mai 1998 (13.05.1998)	US
(30) 60/079,294	25 Mar/mar 1998 (25.03.1998)	US	(30) 60/082,804	22 Apr/avr 1998 (22.04.1998)	US	(30) 60/085,573	15 May/mai 1998 (15.05.1998)	US
(30) 60/079,656	26 Mar/mar 1998 (26.03.1998)	US	(30) 60/082,704	22 Apr/avr 1998 (22.04.1998)	US	(30) 60/085,697	15 May/mai 1998 (15.05.1998)	US
(30) 60/079,728	27 Mar/mar 1998 (27.03.1998)	US	(30) 60/082,767	23 Apr/avr 1998 (23.04.1998)	US	(30) 60/085,580	15 May/mai 1998 (15.05.1998)	US
(30) 60/079,786	27 Mar/mar 1998 (27.03.1998)	US	(30) 60/082,796	23 Apr/avr 1998 (23.04.1998)	US	(30) 60/085,579	15 May/mai 1998 (15.05.1998)	US
(30) 60/079,664	27 Mar/mar 1998 (27.03.1998)	US	(30) 60/083,336	27 Apr/avr 1998 (27.04.1998)	US	(30) 60/085,704	15 May/mai 1998 (15.05.1998)	US
(30) 60/079,689	27 Mar/mar 1998 (27.03.1998)	US	(30) 60/083,322	28 Apr/avr 1998 (28.04.1998)	US	(30) 60/085,582	15 May/mai 1998 (15.05.1998)	US
(30) 60/079,663	27 Mar/mar 1998 (27.03.1998)	US	(30) 60/083,392	29 Apr/avr 1998 (29.04.1998)	US	(30) 60/085,689	15 May/mai 1998 (15.05.1998)	US
(30) 60/079,923	30 Mar/mar 1998 (30.03.1998)	US	(30) 60/083,499	29 Apr/avr 1998 (29.04.1998)	US	(30) 60/085,700	15 May/mai 1998 (15.05.1998)	US
(30) 60/079,920	30 Mar/mar 1998 (30.03.1998)	US	(30) 60/083,545	29 Apr/avr 1998 (29.04.1998)	US	(30) 60/086,023	18 May/mai 1998 (18.05.1998)	US
(30) 60/080,105	31 Mar/mar 1998 (31.03.1998)	US	(30) 60/083,554	29 Apr/avr 1998 (29.04.1998)	US	(30) 60/086,486	22 May/mai 1998 (22.05.1998)	US
(30) 60/080,165	31 Mar/mar 1998 (31.03.1998)	US	(30) 60/083,495	29 Apr/avr 1998 (29.04.1998)	US	(30) 60/086,414	22 May/mai 1998 (22.05.1998)	US
(30) 60/080,194	31 Mar/mar 1998 (31.03.1998)	US	(30) 60/083,558	29 Apr/avr 1998 (29.04.1998)	US	(30) 60/086,392	22 May/mai 1998 (22.05.1998)	US
(30) 60/080,107	31 Mar/mar 1998 (31.03.1998)	US	(30) 60/083,496	29 Apr/avr 1998 (29.04.1998)	US	(30) 60/086,430	22 May/mai 1998 (22.05.1998)	US
(30) 60/080,333	1 Apr/avr 1998 (01.04.1998)	US	(30) 60/083,559	29 Apr/avr 1998 (29.04.1998)	US	(30) 60/087,208	28 May/mai 1998 (28.05.1998)	US
(30) 60/080,327	1 Apr/avr 1998 (01.04.1998)	US	(30) 60/083,500	29 Apr/avr 1998 (29.04.1998)	US	(30) 60/087,098	28 May/mai 1998 (28.05.1998)	US
(30) 60/080,334	1 Apr/avr 1998 (01.04.1998)	US	(30) 60/083,742	30 Apr/avr 1998 (30.04.1998)	US	(30) 60/087,106	28 May/mai 1998 (28.05.1998)	US
(30) 60/080,328	1 Apr/avr 1998 (01.04.1998)	US	(30) 60/084,366	5 May/mai 1998 (05.05.1998)	US	(30) 60/094,651	30 Jul/juil 1998 (30.07.1998)	US
(30) 60/081,071	8 Apr/avr 1998 (08.04.1998)	US	(30) 60/084,441	6 May/mai 1998 (06.05.1998)	US	(30) 60/100,038	11 Sep/sep 1998 (11.09.1998)	US
(30) 60/081,070	8 Apr/avr 1998 (08.04.1998)	US						

deposited with the ATCC. It is understood that the deposited clone contains the actual sequence and that the sequences provided herein are merely representative based on current sequencing techniques. Moreover, given the sequences provided herein and knowledge of the universal genetic code, the corresponding nucleotides for any given amino acid can be routinely identified and vice versa.

Analysis of the amino acid sequence of the full-length PRO273 polypeptide suggests that portions of it possess sequence identity with human macrophage inflammatory protein-2, cytokine-induced neutrophil chemoattractant 2, and neutrophil chemotactic factor 2-beta, thereby indicating that PRO273 is a novel chemokine.

As discussed further below, the cDNA was subcloned into a baculovirus vector and expressed in insect cells as a C-terminally tagged IgG fusion protein. N-terminal sequencing of the resultant protein identified the signal sequence cleavage site, yielding a mature polypeptide of 77 amino acids. The mature sequence, showing 31-40% identity to other human CXC chemokines, includes the four canonical cysteine residues but lacks the ELR motif. Northern analysis demonstrates expression at least in the small intestine, colon, spleen, lymph node and kidney. By in situ hybridization, also described in detail below, mRNA is localized to the lamina propria of intestinal villi and to renal tubules.

EXAMPLE 58: Isolation of cDNA Clones Encoding Human PRO701

A consensus sequence was obtained relative to a variety of EST sequences as described in Example 1 above, wherein the consensus sequence obtained is herein designated DNA39848. Based on the DNA39848 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO701.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-GGCAAGCTACGGAAACGTCATCGTG-3' (SEQ ID NO:376)

reverse PCR primer 5'-AACCCCCGAGCCAAAAGATGGTCAC-3' (SEQ ID NO:377)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA39848 sequence which had the following nucleotide sequence:

hybridization probe

5'-GTACCGGTGACCAGGCAGCAAAAGGCAACTATGGGCTCCTGGATCAG-3' (SEQ ID NO:378).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO701 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue (LIB227).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO701 [herein designated as UNQ365 (DNA44205-1285)] (SEQ ID NO:374) and the derived protein sequence for PRO701.

The entire nucleotide sequence of UNQ365 (DNA44205-1285) is shown in Figure 150 (SEQ ID NO:374). Clone UNQ365 (DNA44205-1285) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 50-52 and ending at the stop codon at nucleotide positions 2498-3000 (Figure 150). The predicted polypeptide precursor is 816 amino acids long (Figure 151). The full-length PRO701 protein shown in Figure 151 has an estimated molecular weight of about 91,794 daltons, a pI of about 5.88 and NX(S/T) being 4. Clone UNQ365 (DNA44205-1285) has been deposited with the ATCC on March 31, 1998. It is understood that the clone was the correct and actual sequence, wherein the sequences provided herein are representative based on

sequencing techniques.

Still regarding the amino acid sequence shown in Figure 151, there is a potential signal peptide cleavage site at about amino acid 25. There are potential N-glycosylation sites at about amino acid positions 83, 511, 716 and 803. The carboxylesterases type-B signature 2 sequence is at about residues 125 to 135. Regions homologous with carboxylesterase type-B are also at about residues 54-74, 197-212 and 221-261. A potential transmembrane region corresponds approximately to amino acids 671 through about 700. The corresponding nucleic acids can be routinely determined from the sequences provided herein.

Analysis of the amino acid sequence of the full-length PRO701 polypeptide suggests that it possess significant homology to the neuroligins from rattus norvegicus indicating that PRO701 may be a novel human neuroligin.

EXAMPLE 59: Isolation of cDNA Clones Encoding Human PRO704

A consensus sequence was obtained relative to a variety of EST sequences as described in Example 1 above, wherein the consensus sequence obtained is herein designated DNA43033. Based on the DNA43033 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO704.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CCTTGGGTCGTGGCAGCAGTGG-3' (SEQ ID NO:381);

reverse PCR primer 5'-CACTCTCCAGGCTGCATGCTCAGG-3' (SEQ ID NO:382).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the DNA43033 consensus sequence which had the following nucleotide sequence:

hybridization probe

5'-GTCAAACGTTCGAGTACTTGAAACGGGAGCACTCGCTGTCGAAGC-3' (SEQ ID NO:383).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO704 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue (LIB227).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO704 [herein designated as UNQ368 (DNA50911-1288)] (SEQ ID NO:379) and the derived protein sequence for PRO704.

The entire nucleotide sequence of UNQ368 (DNA50911-1288) is shown in Figure 152 (SEQ ID NO:379). Clone UNQ368 (DNA50911-1288) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 8-10 and ending at the stop codon at nucleotide positions 1052-1054 (Figure 152). The predicted polypeptide precursor is 348 amino acids long (Figure 153). The full-length PRO704 protein shown in Figure 153 has an estimated molecular weight of about 39,711 and a pI of about 8.7. Clone UNQ368 (DNA50911-1288) has been deposited with the ATCC on March 31, 1998. Regarding the sequence, it is understood that the deposited clone contains the correct sequence, and the sequences provided herein are based on known sequencing techniques.

Analysis of the amino acid sequence of the full-length PRO704 polypeptide suggests that portions of it possess significant homology to the vesicular integral membrane protein 36, thereby indicating that PRO704 may be a novel vesicular integral membrane protein.

Still analyzing the amino acid sequence of SEQ ID NO:380, the putative signal peptide is at about amino acids 1-39 of SEQ ID NO:380. The transmembrane domain is at amino acids 310-335 of SEQ ID NO:380. A potential N-glycosylation site is at about amino acids 180-183 of SEQ ID NO:380. The corresponding nucleotides can be routinely determined given the sequences provided herein.

EXAMPLE 60: Isolation of cDNA Clones Encoding Human PRO706

A consensus sequence was obtained relative to a variety of EST sequences as described in Example 1 above, wherein the consensus sequence obtained is herein designated DNA40669. Based on the DNA40669 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO706.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CCAAGCAGCTTAGAGCTCCAGACC-3' (SEQ ID NO:386)

reverse PCR primer 5'-TTCCCTATGCTCTGTATTGGCATGG-3' (SEQ ID NO:387)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA40669 sequence which had the following nucleotide sequence

hybridization probe

5'-GCCACTTCTGCCACAATGTCAGCTTTCCCTGTACCAGAAATGGCTGTGTT-3' (SEQ ID NO:388)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO706 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal brain tissue (LIB153).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO706 [herein designated as UNQ370 (DNA48329-1290)] (SEQ ID NO:384) and the derived protein sequence for PRO706. It is understood that the deposited clone contains the actual sequence, and that the sequences provided herein are representative based on current sequencing techniques.

The entire nucleotide sequence of UNQ370 (DNA48329-1290) is shown in Figure 154 (SEQ ID NO:384). Clone UNQ370 (DNA48329-1290) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 279-281 and ending at the stop codon at nucleotide positions 1719-1721 (Figure 154). The predicted polypeptide precursor is 480 amino acids long (Figure 155). The full-length PRO706 protein shown in Figure 155 has an estimated molecular weight of about 55,239 daltons and a pI of about 9.30. Clone UNQ370 (DNA48329-1290) has been deposited with the ATCC on April 21, 1998.

Still regarding the amino acid sequence shown in Figure 155, there is a potential signal peptide cleavage site at about amino acid 19. There are potential N-glycosylation sites at about amino acid positions 305 and 354. There is a potential tyrosine kinase phosphorylation site at about amino acid position 333. A region homologous with histidine acid phosphatase is at about residues 87-102. The corresponding nucleic acid regions can be routinely determined given the provided sequences, i.e., the codons can be determined from the specifically named amino acids given.

Analysis of the amino acid sequence of the full-length PRO706 polypeptide suggests that portions of it possess significant homology to the human prostatic acid phosphatase precursor thereby indicating that PRO706 may

be a novel human prostatic acid phosphatase.

EXAMPLE 61: Isolation of cDNA Clones Encoding Human PRO707

A consensus sequence was obtained relative to a variety of EST sequences as described in Example 1 above, wherein the consensus sequence obtained is herein designated DNA42775. Based on DNA42775, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO707.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-TCCGTCTCTGTGAACCGCCCCAC-3' (SEQ ID NO:391);

reverse PCR primer 5'-CTCGGGCGCATTGTCGTTCTGGTC-3' (SEQ ID NO:392).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the DNA42775 sequence which had the following nucleotide sequence:

hybridization probe

5'-CCGACTGTGAAAGAGAACGCCCCAGATCCACTTATTCCCC-3' (SEQ ID NO:393).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO707 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue (LIB227).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO707 [herein designated as UNQ371 (DNA48306-1291)] (SEQ ID NO:389) and the derived protein sequence for PRO707.

The entire nucleotide sequence of UNQ371 (DNA48306-1291) is shown in Figure 156 (SEQ ID NO:389). Clone UNQ371 (DNA48306-1291) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 371-373 and ending at the stop codon at nucleotide positions 3119-3121 of SEQ ID NO:389. The predicted polypeptide precursor is 916 amino acids long (Figure 157). The full-length PRO707 protein shown in Figure 157 has an estimated molecular weight of about 100,204 daltons and a pI of about 4.92. Clone UNQ371 (DNA48306-1291) has been deposited with ATCC on May 27, 1998. It is understood that the clone UNQ371 which is deposited is that which encodes PRO707, and that the sequences herein are merely representations based on known sequencing techniques which may be subject to minor errors.

Regarding analysis of the amino acid sequence, the signal sequence appears to be at about 1 through 30 of SEQ ID NO:390. Cadherins extracellular repeated domain signature sequence is at about amino acids 121-131, 230-240, 335-345, 440-450, and 550-560 of SEQ ID NO:390. Tyrosine kinase phosphorylation sites are at about amino acids 124-132 and 580-586 of SEQ ID NO:390. A potential transmembrane domain is at about amino acids 682-715 ± 5. The nucleic acid positions can be derived by referring to the corresponding codon for the named amino acid.

Analysis of the amino acid sequence of the full-length PRO707 polypeptide suggests that portions of it possess significant homology to the cadherin FIB3 protein, expressed in human fibroblasts, thereby indicating that PRO707 may be a novel cadherin.

EXAMPLE 62: Isolation of cDNA Clones Encoding Human PRO322

A consensus sequence was obtained relative to a variety of EST sequences as described in Example 1 above, wherein the consensus sequence obtained is herein designated DNA48336. Based on the DNA48336 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO322.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CAGCCTACAGAATAAAGATGGCCC-3' (SEQ ID NO:396)

reverse PCR primer 5'-GGTGCAATGATCTGCCAGGCTGAT-3' (SEQ ID NO:397)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the DNA48336 consensus sequence which had the following nucleotide sequence:

hybridization probe

5'-AGAAATACCTGTGGTTCAGTCCATCCCAAACCCCTGCTACAACAGCAG-3' (SEQ ID NO:398).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO322 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue (LIB227).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO322 [herein designated as UNQ283 (DNA48336-1309)] (SEQ ID NO:394) and the derived protein sequence for PRO322. It is understood that UNQ283 (DNA48336-1309) in fact encodes PRO322, and that SEQ ID NO:394 is a representation of the sequence based on sequencing techniques known in the art.

The entire nucleotide sequence of UNQ283 (DNA48336-1309) is shown in Figure 158 (SEQ ID NO:394). Clone UNQ283 (DNA48336-1309) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 166-168 and ending at the stop codon at nucleotide positions 946-948 (Figure 158). The predicted polypeptide precursor is 260 amino acids long (Figure 159). The full-length PRO322 protein shown in Figure 159 has an estimated molecular weight of about 28,028 daltons and a pI of about 7.87. Clone UNQ283 (DNA48336-1309) has been deposited with ATCC and is assigned ATCC deposit no. 209669.

Regarding the amino acid sequence of Figure 159, a potential N-glycosylation site is at amino acid 110 of SEQ ID NO:395. The serine proteases, trypsin family and histidine active site is identified at amino acids 69 through 74 of SEQ ID NO:395 and the consensus sequence is identified at amino acids 207 through 217 of SEQ ID NO:395. The kringle domain proteins motif is identified at amino acids 205 through 217 of SEQ ID NO:395. The putative signal peptide is encoded at about amino acids 1-23.

Analysis of the amino acid sequence of the full-length PRO322 polypeptide suggests that portions of it possess significant homology to neuropsin and other serine proteases, thereby indicating that PRO322 is a novel serine protease related to neuropsin.

EXAMPLE 63: Isolation of cDNA Clones Encoding Human PRO526

A consensus sequence was obtained relative to a variety of EST sequences as described in Example 1 above, wherein the consensus sequence obtained is herein designated DNA39626. Based on the DNA39626 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO526.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-TGGCTGCCCTGCAGTACCTCTACC-3' (SEQ ID NO:401);

reverse PCR primer 5'-CCCTGCAGGTCATTGGCAGCTAGG-3' (SEQ ID NO:402).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the DNA39626 consensus sequence which had the following nucleotide sequence:

hybridization probe

5'-AGGCACTGCCTGATGACACCTTCCGCGACCTGGGCAACCTCACAC-3' (SEQ ID NO:403).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO526 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal liver tissue (LIB228).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO526 [herein designated as UNQ330 (DNA44184-1319)] (SEQ ID NO:399) and the derived protein sequence for PRO526.

The entire nucleotide sequence of UNQ330 (DNA44184-1319) is shown in Figure 160 (SEQ ID NO:399). Clone UNQ330 (DNA44184-1319) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 514-516 and ending at the stop codon at nucleotide positions 1933-1935 (Figure 160). The predicted polypeptide precursor is 473 amino acids long (Figure 161). The full-length PRO526 protein shown in Figure 161 has an estimated molecular weight of about 50,708 daltons and a pI of about 9.28. Clone UNQ330 (DNA44184-1319) has been deposited with the ATCC on March 26, 1998. It is understood that the clone contains the actual sequence, whereas the sequences presented herein are representative based on current sequencing techniques.

Analysis of the amino acid sequence of the full-length PRO526 polypeptide suggests that portions of it possess significant homology to the leucine repeat rich proteins including ALS, SLIT, carboxypeptidase and platelet glycoprotein V thereby indicating that PRO526 is a novel protein which is involved in protein-protein interactions.

Still analyzing SEQ ID NO:400, the signal peptide sequence is at about amino acids 1-26. A leucine zipper pattern is at about amino acids 135-156. A glycosaminoglycan attachment is at about amino acids 436-439. N-glycosylation sites are at about amino acids 82-85, 179-182, 237-240 and 423-426. A von Willebrand factor (VWF) type C domain(s) is found at about amino acids 411-425. The skilled artisan can understand which nucleotides correspond to these amino acids based on the sequences provided herein.

EXAMPLE 64: Isolation of cDNA Clones Encoding Human PRO531

An ECD database was searched and an expressed sequence tag (EST) from LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA was identified which showed homology to protocadherin 3. Based on this sequence, a search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)) as a comparison of the ECD protein sequences to a 6 frame translation of the EST sequence. Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into consensus DNA sequences with the program "phrap" (Phil Green, University of Washington, Seattle, Washington; http://bozeman.mbt.washington.edu/phrap.docs/phrap.html).

A consensus DNA sequence was assembled relative to other EST sequences using phrap. Based on the consensus sequence obtained, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained

the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO531.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CTGAGAACGCGCCTGAAACTGTG-3' (SEQ ID NO:406);

reverse PCR primer 5'-AGCGTTGTCATTGACATCGGCG-3' (SEQ ID NO:407).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA sequence which had the following nucleotide sequence:

hybridization probe

5'-TTAGTTGCTCCATTCAGGAGGATCTACCCTTCCTCCTGAAATCCGCGGAA-3' (SEQ ID NO:408).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO531 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal brain tissue (LIB153). The cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO531 [herein designated as UNQ332 (DNA48314-1320)] (SEQ ID NO:404) and the derived protein sequence for PRO531.

The entire representative nucleotide sequence of UNQ332 (DNA48314-1320) is shown in Figure 162 (SEQ ID NO:404). It is understood that the actual sequence is that within the clone deposited with the ATCC as DNA48314-1320. Clone UNQ332 (DNA48314-1320) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 171-173 and ending at the stop codon at nucleotide positions 2565-2567 (Figure 162). The predicted polypeptide precursor is 789 amino acids long (Figure 163). The full-length PRO531 protein shown in Figure 163 has an estimated molecular weight of about 87,552 daltons and a pI of about 4.84. Clone UNQ332 (DNA48314-1320) has been deposited with the ATCC on March 26, 1998.

Analysis of the amino acid sequence of the full-length PRO531 polypeptide suggests that portions of it possess significant homology to protocadherin 3. Moreover, PRO531 is found in the brain, like other protocadherins, thereby indicating that PRO531 is a novel member of the cadherin superfamily.

Still analyzing the amino acid sequence of SEQ ID NO:405, the cadherin extracellular repeated domain signature is found at about amino acids 122-132, 231-241, 336-346, 439-449 and 549-559 of SEQ ID NO:405. An ATP/GTP-binding site motif A (P-loop) is found at about amino acids 285-292 of SEQ ID NO:405. N-glycosylation sites are found at least at about amino acids 567-570, 786-790, 418-421 and 336-339 of SEQ ID NO:405. The signal peptide is at about amino acids 1-26, and the transmembrane domain is at about amino acids 685-712 of SEQ ID NO:405.

EXAMPLE 65: Isolation of cDNA Clones Encoding Human PRO534

A consensus sequence was obtained relative to a variety of EST sequences as described in Example 1 above, wherein the consensus sequence obtained is herein designated DNA43038. Based on the 43048 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and

2) for use as probes to isolate a clone of the full-length coding sequence for PRO534.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CACAGAGCCAGAAGTGGCGGAATC-3' (SEQ ID NO:411);

reverse PCR primer 5'-CCACATGTTCCTGCTCTTGTCCTGG-3' (SEQ ID NO:412).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA43038 sequence which had the following nucleotide sequence:

hybridization probe

5'-CGGTAGTGACTGTACTCTAGTCCTGTTTTACACCCCGTGGTGCCG-3' (SEQ ID NO:413).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO534 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal lung tissue (LIB26).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO534 [herein designated as UNQ335 (DNA48333-1321)] (SEQ ID NO:409) and the derived protein sequence for PRO534.

The entire nucleotide sequence of UNQ335 (DNA48333-1321) is shown in Figure 164 (SEQ ID NO:409). Clone UNQ335 (DNA48333-1321) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 87-89 and ending at the stop codon at nucleotide positions 1167-1169 (Figure 164). The predicted polypeptide precursor is 360 amino acids long (Figure 165). The full-length PRO534 protein shown in Figure 165 has an estimated molecular weight of about 39,885 daltons and a pI of about 4.79. Clone UNQ335 (DNA48333-1321) has been deposited with ATCC on March 26, 1998. It is understood that the deposited clone contains the actual sequence, and that the sequences provided herein are representative based on current sequencing techniques.

Analysis of the amino acid sequence of the full-length PRO534 polypeptide suggests that portions of it possess significant sequence identity with the protein disulfide isomerase, thereby indicating that PRO534 may be a novel disulfide isomerase.

Still analyzing the amino acid sequence of PRO534, the signal peptides is at about amino acids 1-25 of SEQ ID NO:410. The transmembrane domain is at about amino acids 321-340 of SEQ ID NO:410. The disulfide isomerase corresponding region is at amino acids 212-302 of SEQ ID NO:410. The thioredoxin domain is at amino acids 211-227 of SEQ ID NO:410. N-glycosylation sites are at: 165-168, 181-184, 187-190, 194-197, 206-209, 278-281, and 293-296 of SEQ ID NO:410. The corresponding nucleotides can routinely be determined from the sequences provided herein. PRO534 has a transmembrane domain rather than an ER retention peptide like other protein disulfide isomerases. Additionally, PRO534 may have an intron at the 5 prime end.

EXAMPLE 66: Isolation of cDNA Clones Encoding Human PRO697

A consensus sequence was obtained relative to a variety of EST sequences as described in Example 1 above, wherein the consensus sequence obtained is herein designated DNA43052. Based on this consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO697.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CCTGGCTCGCTGCTGCTGCTC-3' (SEQ ID NO:416);

reverse PCR primer 5'-CCTCACAGGTGCACTGCAAGCTGTC-3' (SEQ ID NO:417).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the DNA43052 consensus sequence which had the following nucleotide sequence:

hybridization probe

5'-CTCTTCCTCTTTGGCCAGCCCGACTTCTCCTACAAGCGCAGAATTGC-3' (SEQ ID NO:418).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO697 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue (LIB227).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO697 [herein designated as UNQ361 (DNA50920-1325)] (SEQ ID NO:414) and the derived protein sequence for PRO697.

The entire nucleotide sequence of UNQ361 (DNA50920-1325) is shown in Figure 166 (SEQ ID NO:414). Clone UNQ361 (DNA50920-1325) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 44-46 and ending at the stop codon at nucleotide positions 929-931 (Figure 166). The predicted polypeptide precursor is 295 amino acids long (Figure 167). The full-length PRO697 protein shown in Figure 167 has an estimated molecular weight of about 33,518 daltons and a pI of about 7.74. Clone UNQ361 (DNA50920-1325) was deposited with the ATCC on March 26, 1998. It is understood that the deposited clone contains the actual sequence, and that the sequences provided herein are representative based on current sequencing techniques.

Analysis of the amino acid sequence of the full-length PRO697 polypeptide suggests that portions of it possess significant sequence identity with sFRPs, thereby indicating that PRO697 may be a novel sFRP family member.

Still analyzing the amino acid sequence of PRO697, the signal peptides is at about amino acids 1-20 of SEQ ID NO:415. The cystein rich domain, having identity with the frizzled N-terminus, is at about amino acids 6-153 of SEQ ID NO:415. The corresponding nucleotides can routinely be determined from the sequences provided herein.

EXAMPLE 67: Isolation of cDNA Clones Encoding Human PRO717

A consensus sequence was obtained relative to a variety of EST sequences as described in Example 1 above, wherein the consensus sequence obtained is herein designated DNA42829. Based on the DNA42829 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO717.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-AGCTTCTCAGCCCTCCTGGAGCAG-3' (SEQ ID NO:421);

reverse PCR primer 5'-CGGGTCAATAAACCTGGACGCTTGG-3' (SEQ ID NO:422).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the DNA42829 consensus sequence which had the following nucleotide sequence:

hybridization probe

5'-TATGTGGACCGGACCAAGCACTTCACTGAGGCCACCAAGATTG-3' (SEQ ID NO:423).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones

encoding the PRO717 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal liver tissue (LIB229).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO717 [herein designated as UNQ385 (DNA50988-1326)] (SEQ ID NO:419) and the derived protein sequence for PRO717.

The entire nucleotide sequence of UNQ385 (DNA50988-1326) is shown in Figure 168 (SEQ ID NO:419). Clone UNQ385 (DNA50988-1326) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 17-19 and ending at the stop codon at nucleotide positions 1697-1699 (Figure 168). The predicted polypeptide precursor is 560 amino acids long (Figure 169). The full-length PRO717 protein shown in Figure 169 has an estimated molecular weight of about 58,427 daltons and a pI of about 6.86. Clone UNQ385 (DNA50988-1326) has been deposited with the ATCC on April 28, 1998. Regarding the sequence, it is understood that the deposited clone contains the correct sequence, and the sequences provided herein are based on known sequencing techniques.

Analysis of the amino acid sequence of the full-length PRO717 polypeptide suggests that PRO717 may be a novel 12 transmembrane receptor. The reverse complement strand of DNA50988 has a stretch that matches identically with human regulatory myosin light strand.

Still analyzing the amino acid sequence of SEQ ID NO:420, transmembrane domains are at about amino acids 30-50, 61-79, 98-112, 126-146, 169-182, 201-215, 248-268, 280-300, 318-337, 341-357, 375-387, and 420-441 of SEQ ID NO:420. N-glycosylation sites are at about amino acids 40-43 and 43-46 of SEQ ID NO:420. A glycosaminoglycan attachment site is at about amino acids 468-471 of SEQ ID NO:420. The corresponding nucleotides can be routinely determined given the sequences provided herein.

EXAMPLE 68: Isolation of cDNA Clones Encoding Human PRO731

A database was used to search expressed sequence tag (EST) databases. The EST database used herein was the proprietary EST DNA database LIFESEQ™, of Incyte Pharmaceuticals, Palo Alto, CA. Incyte clone 2581326 was herein identified and termed DNA42801. Based on the DNA42801 sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO731.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-GTAAGCACATGCCTCCAGAGGTGC-3' (SEQ ID NO:426);

reverse PCR primer 5'-GTGACGTGGATGCTTGGGATGTTG-3' (SEQ ID NO:427).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the DNA42801 sequence which had the following nucleotide sequence:

hybridization probe

5'-TGGACACCTTCAGTATTGATGCCAAGACAGGCCAGGTCATTCTGCGTCGA-3' (SEQ ID NO:428).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO731 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human bone marrow tissue (LIB255). The cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI

hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO731 [herein designated as UNQ395 (DNA48331-1329)] (SEQ ID NO:424) and the derived protein sequence for PRO731.

The entire nucleotide sequence of UNQ395 (DNA48331-1329) is shown in Figures 170A-B (SEQ ID NO:424). Clone UNQ395 (DNA48331-1329) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 329-331 and ending at the stop codon at nucleotide positions 3881-3883 (Figures 170A-B). The predicted polypeptide precursor is 1184 amino acids long (Figure 171). The full-length PRO731 protein shown in Figure 171 has an estimated molecular weight of about 129,022 daltons and a pI of about 5.2. Clone UNQ395 (DNA48331-1329) was deposited with the ATCC on March 31, 1998. Regarding the sequence, it is understood that the deposited clone contains the correct sequence, and the sequences provided herein are based on known sequencing techniques.

Analysis of the amino acid sequence of the full-length PRO731 polypeptide suggests that portions of it possess significant identity and similarity to members of the protocadherin family, thereby indicating that PRO731 may be a novel protocadherin.

Still analyzing the amino acid sequence of SEQ ID NO:425, the putative signal peptide is at about amino acids 1-13 of SEQ ID NO:425. The transmembrane domain is at amino acids 719-739 of SEQ ID NO:425. The N-glycosylation of SEQ ID NO:425 are as follows: 415-418, 582-586, 659-662, 662-665, and 857-860. The cadherin extracellular repeated domain signatures are at about amino acids (of SEQ ID NO:425): 123-133, 232-242, 340-350, 448-458, and 553-563. The corresponding nucleotides can be routinely determined given the sequences provided herein.

EXAMPLE 69: Isolation of cDNA Clones Encoding Human PRO218

A consensus sequence was obtained relative to a variety of EST sequences as described in Example 1 above, wherein the consensus sequence obtained is herein designated DNA17411. Two proprietary Genentech EST sequences were employed in the consensus assembly and are shown in Figure 174 and 175. Based on the DNA17411 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO218.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-AAGTGGAGCCGGAGCCTTCC-3' (SEQ ID NO:433);

reverse PCR primer 5'-TCGTTGTTTATGCAGTAGTCGG-3' (SEQ ID NO:434).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA17411 sequence which had the following nucleotide sequence:

hybridization probe

5'-ATTGTTTAAAGACTATGAGATACGTCAGTATGTTGTACAGG-3' (SEQ ID NO:435).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO218 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of

the cDNA libraries was isolated from human fetal kidney tissue (LIB28).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO218 [herein designated as UNQ192 (DNA30867-1335)] (SEQ ID NO:429) and the derived protein sequence for PRO218.

The entire nucleotide sequence of UNQ192 (DNA30867-1335) is shown in Figure 172 (SEQ ID NO:429). Clone UNQ192 (DNA30867-1335) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 150-152 and ending at the stop codon at nucleotide positions 1515-1517 (Figure 172). The predicted polypeptide precursor is 455 amino acids long (Figure 173). The full-length PRO218 protein shown in Figure 173 has an estimated molecular weight of about 52,917 daltons and a pI of about 9.5. Clone UNQ192 (DNA30867-1335) has been deposited with the ATCC on April 28, 1998. Regarding the sequence, it is understood that the deposited clone contains the correct sequence, and the sequences provided herein are based on known sequencing techniques.

Analysis of the amino acid sequence of the full-length PRO218 polypeptide suggests that PRO218 may be a novel transmembrane protein.

Still analyzing the amino acid sequence of SEQ ID NO:430, the putative signal peptide is at about amino acids 1 through 23 of SEQ ID NO:430. Transmembrane domains are potentially at about amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398, and 425-444 of SEQ ID NO:430. N-glycosylation sites are at about amino acids 67, 180, and 243 of SEQ ID NO:430. Eukaryotic cobalamin-binding protein is at about amino acids 151-160 of SEQ ID NO:430. The corresponding nucleotides can be routinely determined given the sequences provided herein.

EXAMPLE 70: Isolation of cDNA Clones Encoding Human PRO768

A consensus sequence was obtained relative to a variety of EST sequences as described in Example 1 above, wherein the consensus sequence obtained is herein designated DNA43448. Based on the DNA43448 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO768.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-GGCTGACACCGCAGTGCTCTTCAG-3' (SEQ ID NO:438);

reverse PCR primer 5'-GCTGCTGGGGACTGCAATGTAGCTG-3' (SEQ ID NO:439).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the DNA43448 consensus sequence which had the following nucleotide sequence:

hybridization probe

5'-CATCCTCCATGTCTCCCATGAGGTCTCTATTGCTCCACGAAGCATC-3' (SEQ ID NO:440).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO768 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human bone marrow tissue (LIB255).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO768 [herein designated as UNQ406 (DNA55737-1345)] (SEQ ID NO:436) and the derived protein sequence for PRO768.

The entire nucleotide sequence of UNQ406 (DNA55737-1345) is shown in Figures 176A-B (SEQ ID NO:436). Clone UNQ406 (DNA55737-1345) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 20-22 and ending at the stop codon at nucleotide positions 3443-3445 (Figures

176A-B). The predicted polypeptide precursor is 1141 amino acids long (Figure 177). The full-length PRO768 protein shown in Figure 177 has an estimated molecular weight of about 124,671 daltons and a pI of about 5.82. Clone UNQ406 (DNA55737-1345) has been deposited with the ATCC on April 6, 1998. Regarding the sequence, it is understood that the deposited clone contains the correct sequence, and the sequences provided herein are based on known sequencing techniques.

Analysis of the amino acid sequence of the full-length PRO768 polypeptide suggests that portions of it possess significant sequence identity and similarity with integrin 7.

Still analyzing the amino acid sequence of SEQ ID NO:437, the putative signal peptide is at about amino acids 1-33 of SEQ ID NO:437. The transmembrane domain is at amino acids 1039-1064 of SEQ ID NO:437. N-glycosylation sites are at amino acids: 86-89, 746-749, 949-952, 985-988 and 1005-1008 of SEQ ID NO:437. Integrin alpha chain protein domains are identified at about amino acids: 1064-1071, 384-409, 1041-1071, 317-346, 443-465, 385-407, 215-224, 634-647, 85-99, 322-346, 470-479, 442-466, 379-408 and 1031-1047 of SEQ ID NO:437. The corresponding nucleotides can be routinely determined given the sequences provided herein.

EXAMPLE 71: Isolation of cDNA Clones Encoding Human PRO771

A consensus sequence was obtained relative to a variety of EST sequences as described in Example 1 above, wherein the consensus sequence obtained is herein designated DNA43330. Based on the DNA43330 sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO771.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CAGCAATATTCAGAAGCGGCAAGGG-3' (SEQ ID NO:443);

reverse PCR primer 5'-CATCATGGTCATCACCACCATCATCATC-3' (SEQ ID NO:444).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the DNA43330 consensus sequence which had the following nucleotide sequence:

hybridization probe

5'-GGTTACTACAAGCCAACACAATGTCATGGCAGTGTTGGACAGTGCTGG-3' (SEQ ID NO:445).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO771 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue (LIB28).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO771 [herein designated as UNQ409 (DNA49829-1346)] (SEQ ID NO:441) and the derived protein sequence for PRO771.

The entire nucleotide sequence of UNQ409 (DNA49829-1346) is shown in Figure 178 (SEQ ID NO:441). Clone UNQ409 (DNA49829-1346) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 134-136 and ending at the stop codon at nucleotide positions 1442-1444 (Figure 178). The predicted polypeptide precursor is 436 amino acids long (Figure 179). The full-length PRO771 protein shown in Figure 179 has an estimated molecular weight of about 49,429 daltons and a pI of about 4.8. Clone UNQ409 (DNA49829-1346) has been deposited with the ATCC on April 7, 1998. Regarding the sequence, it is understood that the deposited clone contains the correct sequence, and the sequences provided herein are based on known sequencing techniques.

Analysis of the amino acid sequence of the full-length PRO771 polypeptide suggests that portions of it possess significant homology to the testican protein, thereby indicating that PRO771 may be a novel testican homologue.

Still analyzing the amino acid sequence of SEQ ID NO:442, the putative signal peptide, leucine zipper pattern, N-myristoylation sites, and thyroglobulin type-1 repeats are also shown in Figure 179. The corresponding nucleotides can be routinely determined given the sequences provided herein.

EXAMPLE 72: Isolation of cDNA Clones Encoding Human PRO733

A consensus sequence was obtained relative to a variety of EST sequences as described in Example 1 above, wherein the consensus sequence obtained is herein designated DNA45600. Based on the DNA45600 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO733.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CCCAGCAGGGATGGGCGACAAGA-3' (SEQ ID NO:448);

reverse PCR primer 5'-GTCTTCCAGTTTCATATCCAATA-3' (SEQ ID NO:449).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the DNA45600 consensus sequence which had the following nucleotide sequence:

hybridization probe

5'-CCAGAAGGAGCACGGGGAAGGGCAGCCAGATCTTGTCGCCCAT-3' (SEQ ID NO:450).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO733 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human bone marrow tissue (LIB255).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO733 [herein designated as UNQ411 (DNA52196-1348)] (SEQ ID NO:446) and the derived protein sequence for PRO733.

The entire nucleotide sequence of UNQ411 (DNA52196-1348) is shown in Figures 180A-B (SEQ ID NO:446). Clone UNQ411 (DNA52196-1348) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 106-108 and ending at the stop codon at nucleotide positions 793-795 (Figures 180A-B). The predicted polypeptide precursor is 229 amino acids long (Figure 181). The full-length PRO733 protein shown in Figure 181 has an estimated molecular weight of about 26,017 daltons and a pI of about 4.73. Clone UNQ411 (DNA52196-1348) has been deposited with the ATCC on April 7, 1998. Regarding the sequence, it is understood that the deposited clone contains the correct sequence, and the sequences provided herein are based on known sequencing techniques.

Analysis of the amino acid sequence of the full-length PRO733 polypeptide suggests that portions of it possess significant sequence identity and similarity to the T1/ST2 receptor binding protein precursor and therefore may have a similar function in cell signaling. If it is a cytokine, it may be useful in the treatment of inflammation and cancer.

Still analyzing the amino acid sequence of SEQ ID NO:447, the putative signal peptide, transmembrane domain, N-myristoylation site, and tyrosine kinase site are also shown in Figure 181. The corresponding nucleotides can be routinely determined given the sequences provided herein.

EXAMPLE 73: Isolation of cDNA Clones Encoding Human PRO162

An expressed sequence tag (EST) DNA database (Merck/Washington University) was searched and an EST AA397543 was identified which showed homology to human pancreatitis-associated protein. The EST AA397543 cole was purchased and its insert obtained and sequenced and the sequence obtained is shown in Figure 182 (SEQ ID NO:451).

The entire nucleotide sequence of PRO162 is shown in Figure 182 (SEQ ID NO:451). DNA sequencing of the clone gave the full-length DNA sequence for PRO162 [herein designated as UNQ429 (DNA56965-1356)] (SEQ ID NO:451) and the derived protein sequence for PRO162. Clone UNQ429 (DNA56965-1356) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 86-88 and ending at the stop codon at nucleotide positions 611-613 (Figure 182). The predicted polypeptide precursor is 175 amino acids long (Figure 183). The full-length PRO162 protein shown in Figure 183 has an estimated molecular weight of about 19,330 daltons and a pI of about 7.25. Clone UNQ429 (DNA56965-1356) has been deposited with the ATCC. Regarding the sequence, it is understood that the deposited clone contains the correct sequence, and the sequences provided herein are based on known sequencing techniques.

Analysis of the amino acid sequence of the full-length PRO162 polypeptide suggests that portions of it possess significant homology to the human pancreatitis-associated protein, thereby indicating that PRO162 may be a novel pancreatitis-associated protein.

Still analyzing the amino acid sequence of SEQ ID NO:452, the putative signal peptide is at about amino acids 1-26 of SEQ ID NO:452. A C-type lectin domain signature is at about amino acids 146-171 of SEQ ID NO:452. The corresponding nucleotides can be routinely determined given the sequences provided herein.

EXAMPLE 74: Isolation of cDNA Clones Encoding Human PRO788

A consensus DNA sequence (designated herein as DNA49308) was assembled relative to other EST sequences using phrap as described in Example 1 above. Based upon an observed homology between the DNA49308 consensus sequence and the Incyte EST cloOne no. 2777282, the Incyte EST clone no. 2777282 was purchased and its insert obtained and sequenced, which gave the full-length DNA sequence for PRO788 [herein designated as UNQ430 (DNA56405-1357)] (SEQ ID NO:453) and the derived protein sequence for PRO788.

Clone UNQ430 (DNA56405-1357) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 84-86 and ending at the stop codon at nucleotide positions 459-461 (Figure 184). The predicted polypeptide precursor is 125 amino acids long (Figure 185). The full-length PRO788 protein shown in Figure 185 has an estimated molecular weight of about 13,115 daltons and a pI of about 5.90. Clone UNQ430 (DNA56405-1357) has been deposited with the ATCC. Regarding the sequence, it is understood that the deposited clone contains the correct sequence, and the sequences provided herein are based on known sequencing techniques.

Still analyzing Figure 185, a signal peptide is shown at about amino acids 1-17 of SEQ ID NO:454. An N-glycosylation site is at about amino acids 46-49 of SEQ ID NO:454.

EXAMPLE 75: Isolation of cDNA Clones Encoding Human PRO1008

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated as DNA49804. An EST proprietary to Genentech was employed in the consensus assembly and is herein designated as DNA16508 (Figure 188; SEQ ID NO:457).

Based upon an observed homology between the DNA49804 sequence and Merck EST clone no. AA143670, the Merck EST clone no. AA143670 was purchased and its insert obtained and sequenced. That sequence is shown herein in Figure 186 (SEQ ID NO:455).

Sequencing gave the full length sequence for PRO1008 [herein designated as UNQ492 (DNA57530-1375)] (SEQ ID NO:455) and the derived protein sequence for PRO1008 were identified.

The entire nucleotide sequence of UNQ492 (DNA57530-1375) is shown in Figure 186 (SEQ ID NO:455). Clone UNQ492 (DNA57530-1375) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 138-140 and ending at the stop codon at nucleotide positions 936-938 (Figure 186). The predicted polypeptide precursor is 266 amino acids long (Figure 187). The full-length PRO1008 protein shown in Figure 187 has an estimated molecular weight of about 28,672 daltons and a pI of about 8.85. Clone UNQ492 (DNA57530-1375) has been deposited with the ATCC on May 20, 1998. Regarding the sequence, it is understood that the deposited clone contains the correct sequence, and the sequences provided herein are based on known sequencing techniques.

Analysis of the amino acid sequence of the full-length PRO1008 polypeptide suggests that portions of it possess significant sequence identity and/or similarity with mdkk-1, thereby indicating that PRO1008 may be a novel member of this family and have head inducing activity.

Still analyzing the amino acid sequence of SEQ ID NO:456, the putative signal peptide is at about amino acids 1-23 of SEQ ID NO:456. The N-glycosylation site is at about amino acids 256-259 of SEQ ID NO:456, and the fungal zn-(2)-cys(6) binuclear cluster domain is at about amino acids 110-126 of SEQ ID NO:456. The corresponding nucleotides can of all the amino acids can be routinely determined given the sequences provided herein.

EXAMPLE 76: Isolation of cDNA Clones Encoding Human PRO1012

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above, wherein the consensus sequence is herein designated DNA49313. Based on the DNA49313 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1012.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-ACTCCCCAGGCTGTTCACACTGCC-3' (SEQ ID NO:460);

reverse PCR primer 5'-GATCAGCCAGCCAATACCAGCAGC-3' (SEQ ID NO:461).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the DNA49313 consensus sequence which had the following nucleotide sequence:

hybridization probe

5'-GTGGTGATGATAGAATGCTTTGCCGAATGAAAGGAGTCAACAGCTATCCC-3' (SEQ ID NO:462).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO1012 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue (LIB227).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1012 [herein designated as UNQ495 (DNA56439-1376)] (SEQ ID NO:458) and the derived protein sequence for PRO1012.

The entire nucleotide sequence of UNQ495 (DNA56439-1376) is shown in Figures 189A-B (SEQ ID NO:458). Clone UNQ495 (DNA56439-1376) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 404-406 and ending at the stop codon at nucleotide positions 2645-2647 (Figures 189A-B). The predicted polypeptide precursor is 747 amino acids long (Figure 190). The full-length PRO1012 protein shown in Figure 190 has an estimated molecular weight of about 86,127 daltons and a pI of about 7.46. Clone UNQ495 (DNA56439-1376) has been deposited with ATCC on May 14, 1998. Regarding the sequence, it is understood that the deposited clone contains the correct sequence, and the sequences provided herein are based on known sequencing techniques.

Analysis of the amino acid sequence of the full-length PRO1012 polypeptide suggests that portions of it possess sequence identity with disulfide isomerase thereby indicating that PRO1012 may be a novel disulfide isomerase related protein.

Still analyzing the amino acid sequence of SEQ ID NO:459, the cytochrome C family heme-binding site signature is at about amino acids 158-163 of SEQ ID NO:459. The Nt-DNAJ domain signature is at about amino acids 77-96 of SEQ ID NO:459. An N-glycosylation site is at about amino acids 484-487 of SEQ ID NO:459. The ER targeting sequence is at about amino acids 744-747 of SEQ ID NO:459. It is understood that the polypeptide and nucleic acids disclosed can be routinely formed with or without, these portions as desired, in alternative embodiments. For example, it may be desirable to produce PRO1012 without the ER targeting sequence. The corresponding nucleotides can be routinely determined given the sequences provided herein.

EXAMPLE 77: Isolation of cDNA Clones Encoding Human PRO1014

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 abobe, wherein the consensus sequence obtained is herein designated DNA49811. Based upon an observed homology between the DNA49811 sequence and Incyte EST clone no. 2612207, Incyte EST clone no. 2612207 was purchased and its insert was obtained and sequenced, wherein the sequence obtained is shown in Figure 191 (SEQ OD NO:463).

DNA sequencing gave the full-length DNA sequence for PRO1014 [herein designated as UNQ497 (DNA56409-1377)] (SEQ ID NO:463) and the derived protein sequence for PRO1014.

The entire nucleotide sequence of UNQ497 (DNA56409-1377) is shown in Figure 191 (SEQ ID NO:463). Clone UNQ497 (DNA56409-1377) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 66-68 and ending at the stop codon at nucleotide positions 966-968 (Figure 191). The predicted polypeptide precursor is 300 amino acids long (Figure 192). The full-length PRO1014 protein shown in Figure 192 has an estimated molecular weight of about 33,655 daltons and a pI of about 9.31. Clone UNQ497 (DNA56409-1377) has been deposited with the ATCC on May 20, 1998. Regarding the sequence, it is understood that the deposited clone contains the correct sequence, and the sequences provided herein are based on known sequencing techniques.

Analysis of the amino acid sequence of the full-length PRO1014 polypeptide suggests that portions of it possess sequence identity with reductase, thereby indicating that PRO1014 may be a novel member of the reductase family.

Still analyzing the amino acid sequence of SEQ ID NO:464, the putative signal peptide is at about amino acids 1-19 of SEQ ID NO:464. The cAMP and cGMP dependent protein kinase phosphorylation sites are at about

amino acids 30-33 and 58-61 of SEQ ID NO:464. Short chain alcohol dehydrogenase family proteins are at about amino acids 165-202, 37-49, 112-122 and 210-219 of SEQ ID NO:464. The corresponding nucleotides of these domains and any other amino acids provided herein can be routinely determined given the sequences provided herein.

EXAMPLE 78: Isolation of cDNA Clones Encoding Human PRO1017

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above, wherein that consensus DNA sequence is herein designated DNA53235. Based upon an observed homology between the DNA53235 consensus sequence and the Merck EST clone no. AA243086, the Merck EST clone no. AA243086 was purchased and its insert obtained and sequenced, wherein the sequence obtained is shown in Figure 193 (SEQ ID NO:465). DNA sequencing gave the full-length DNA sequence for PRO1017 [herein designated as UNQ500 (DNA56112-1379)] (SEQ ID NO:465) and the derived protein sequence for PRO1017.

The entire nucleotide sequence of UNQ500 (DNA56112-1379) is shown in Figure 193 (SEQ ID NO:465). Clone UNQ500 (DNA56112-1379) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 128-130 and ending at the stop codon at nucleotide positions 1370-1372 (Figure 193). The predicted polypeptide precursor is 414 amino acids long (Figure 194). The full-length PRO1017 protein shown in Figure 194 has an estimated molecular weight of about 48,414 daltons and a pI of about 9.54. Clone UNQ500 (DNA56112-1379) has been deposited with the ATCC. Regarding the sequence, it is understood that the deposited clone contains the correct sequence, and the sequences provided herein are based on known sequencing techniques.

Analysis of the amino acid sequence of the full-length PRO1017 polypeptide suggests that portions of it possess sequence identity with HNK-1 sulfotransferase, thereby indicating that PRO1017 may be a novel sulfotransferase.

Still analyzing the amino acid sequence of SEQ ID NO:466, the putative signal peptide is at about amino acids 1-31 of SEQ ID NO:466. N-glycosylation sites are at about amino acids 134-137, 209-212, 280-283 and 370-273 of SEQ ID NO:466. The TNFR/NGFR family cystein-rich region protein is at about amino acids 329-332 of SEQ ID NO:466. The corresponding nucleotides can be routinely determined given the sequences provided herein. The protein can be secreted.

EXAMPLE 79: Isolation of cDNA Clones Encoding Human PRO474

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above, wherein the consensus sequence obtained is herein designated DNA49818. Based upon an observed homology between the DNA49818 consensus sequence and the Merck EST clone no. H77889, the Merck EST clone no. H77889 was purchased and its insert obtained and sequenced, wherein the sequence obtained is herein shown in Figure 195 (SEQ ID NO:467). DNA sequencing gave the full-length DNA sequence for PRO474 [herein designated as UNQ502 (DNA56045-1380)] (SEQ ID NO:467) and the derived protein sequence for PRO474.

The entire nucleotide sequence of UNQ502 (DNA56045-1380) is shown in Figure 195 (SEQ ID NO:467). Clone UNQ502 (DNA56045-1380) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 106-108 and ending at the stop codon at nucleotide positions 916-918 (Figure 195). The predicted polypeptide precursor is 270 amino acids long (Figure 196). The full-length PRO474 protein shown in Figure 196 has an estimated molecular weight of about 28,317 daltons and a pI of about 6.0. Clone UNQ502

(DNA56045-1380) has been deposited with the ATCC. Regarding the sequence, it is understood that the deposited clone contains the correct sequence, and the sequences provided herein are based on known sequencing techniques.

Still analyzing the amino acid sequence of SEQ ID NO:468, an N-glycosylation site is at about amino acids 138-141 of SEQ ID NO:468. Short-chain alcohol dehydrogenase family proteins are at about amino acids 10-22, 81-91, 134-171 and 176-185 of SEQ ID NO:468. The corresponding nucleotides can be routinely determined given the sequences provided herein.

EXAMPLE 80: Isolation of cDNA Clones Encoding Human PRO1031

An initial consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above, wherein the consensus sequence obtained is herein designated as DNA47332. Based upon an observed homology between the DNA47332 sequence and the Merck EST clone no. W74558, Merck EST clone no. W74558 was purchased and its insert obtained and sequenced, wherein the sequence obtained is shown in Figure 197 (SEQ ID NO:469). DNA sequencing gave the full-length DNA sequence for PRO1031 [herein designated as UNQ516 (DNA59294-1381)] (SEQ ID NO:469) and the derived protein sequence for PRO1031.

The entire nucleotide sequence of UNQ516 (DNA59294-1381) is shown in Figure 197 (SEQ ID NO:469). Clone UNQ516 (DNA59294-1381) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 42-44 and ending at the stop codon at nucleotide positions 582-584 (Figure 197). The predicted polypeptide precursor is 180 amino acids long (Figure 198). The full-length PRO1031 protein shown in Figure 198 has an estimated molecular weight of about 20,437 daltons and a pI of about 9.58. Clone UNQ516 (DNA59294-1381) has been deposited with the ATCC. Regarding the sequence, it is understood that the deposited clone contains the correct sequence, and the sequences provided herein are based on known sequencing techniques.

Analysis of the amino acid sequence of the full-length PRO1031 polypeptide suggests that it is a novel cytokine.

Still analyzing the amino acid sequence of SEQ ID NO:470, the putative signal peptide is at about amino acids 1-20 of SEQ ID NO:470. An N-glycosylation site is at about amino acids 75-78 of SEQ ID NO:470. A region having sequence identity with IL-17 is at about amino acids 96-180. The corresponding nucleotides can be routinely determined given the sequences provided herein.

EXAMPLE 81: Isolation of cDNA Clones Encoding Human PRO938

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above, wherein that consensus sequence is herein designated DNA49798. Based on the DNA49798 DNA consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO938.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-GTCCAGCCCATGACCGCCTCCAAC-3' (SEQ ID NO:473)

reverse PCR primer 5'-CTCTCCTCATCCACACCAGCAGCC-3' (SEQ ID NO:474)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA49798 sequence which had the following nucleotide sequence:

hybridization probe

5'-GTGGATGCTGAAATTTTACGCCCCATGGTGTCCATCCTGCCAGC-3' (SEQ ID NO:475)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO938 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue (LIB227).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO938 [herein designated as UNQ475 (DNA56433-1406)] (SEQ ID NO:471) and the derived protein sequence for PRO938.

The entire nucleotide sequence of UNQ475 (DNA56433-1406) is shown in Figure 199 (SEQ ID NO:471). Clone UNQ475 (DNA56433-1406) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 134-136 and ending at the stop codon at nucleotide positions 1181-1183 (Figure 199). The predicted polypeptide precursor is 349 amino acids long (Figure 200). The full-length PRO938 protein shown in Figure 200 has an estimated molecular weight of about 38,952 daltons and a pI of about 4.34. Analysis of the full-length PRO938 sequence shown in Figure 200 (SEQ ID NO:472) evidences the presence of the following features: a signal peptide from amino 1 to about amino acid 22, a transmembrane domain from about amino acid 191 to about amino acid 211, a potential N-glycosylation site from about amino acid 46 to about amino acid 49, a region homologous to disulfide isomerase from about amino acid 56 to about amino acid 72, and a region having sequence identity with flavodoxin proteins from about amino acid 173 to about amino acid 187.

Clone UNQ475 (DNA56433-1406) has been deposited with ATCC on May 12, 1998, and is assigned ATCC Accession No. 209857.

Analysis of the amino acid sequence of the full-length PRO938 polypeptide suggests that it possesses significant sequence similarity to protein disulfide isomerase, thereby indicating that PRO938 may be a novel protein disulfide isomerase. An analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced significant homology between the PRO938 amino acid sequence and the following Dayhoff sequences, P_W03626, P_W03627, P_R70491, GARP_PLAFF, XLU85970_1, ACADISPROA_1, IE68_HSVSA, KSU52064_1, U93872_83, P_R97866.

EXAMPLE 82: Isolation of cDNA Clones Encoding Human PRO1082

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above, wheein the consensus sequence is herein designated DNA38097. Based on this consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1082.

A set of PCR primers (two forward and one reverse) were synthesized:

forward primer 1 5'-GTCCACAGACAGTCATCTCAGGAGCAG-3' (SEQ ID NO:478);

forward primer 2 5'-ACAAGTGTCTTCCCAACCTG-3' (SEQ ID NO:479);

reverse primer 1 5'-ATCCTCCCAGAGCCATGGTACCTC-3' (SEQ ID NO:480).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the DNA38097 consensus sequence which had the following nucleotide sequence:

hybridization probe

5'-CCAAGGATAGCTGTTGTTTCAGAGAAAGGATCGTGTGCTGCATCTCCTCCT-3' (SEQ ID NO:481).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primers identified above. A positive library was then used to isolate clones encoding the PRO1082 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue (LIB227).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1082 [herein designated as UNQ539 (DNA53912-1457)] (SEQ ID NO:476) and the derived protein sequence for PRO1082.

The entire nucleotide sequence of UNQ539 (DNA53912-1457) is shown in Figure 201 (SEQ ID NO:476). Clone UNQ539 (DNA53912-1457) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 160-162 and ending at the stop codon at nucleotide positions 763-765 (Figure 201). The predicted polypeptide precursor is 201 amino acids long (Figure 202). The full-length PRO1082 protein shown in Figure 202 has an estimated molecular weight of about 22,563 daltons and a pI of about 4.87. Clone UNQ539 (DNA53912-1457) has been deposited with the ATCC. Regarding the sequence, it is understood that the deposited clone contains the correct sequence, and the sequences provided herein are based on known sequencing techniques.

Still analyzing the amino acid sequence of SEQ ID NO:477, the transmembrane domain is at about amino acids 45-65 of SEQ ID NO:477. A cAMP- and cGMP-dependent protein kinase phosphorylation site is at about amino acids 197-200 of SEQ ID NO:477. N-myristoylation sites are at about amino acids 35-40 and 151-156 of SEQ ID NO:477. The regions which share sequence identity with the LDL receptor are at about amino acids 34-67 and 70-200 of SEQ ID NO:477. The corresponding nucleotides of these amino acid regions and others can be routinely determined given the sequences provided herein.

EXAMPLE 83: Isolation of cDNA Clones Encoding Human PRO1083

A cDNA sequence was identified using the amylase screening technique described in Example 2 above, wherein that cDNA sequence is designated herein as DNA24256 (Figure 205; SEQ ID NO:484). That cDNA sequence was then compared and aligned with other known EST sequencees as described in Example 1 above to obtain a consensus DNA sequence which is designated herein as DNA43422. Based on the DNA 43422 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1083.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-GGCATTGGAGCAGTGCTGGGTG-3' (SEQ ID NO:485);

reverse PCR primer 5'-TGGAGGCCTAGATGCGGCTGGACG-3' (SEQ ID NO:486).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO1083 gene using the reverse PCR primer. RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue (LIB227).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1083 [herein designated as UNQ540 (DNA50921-1458)] (SEQ ID NO:482) and the derived protein sequence for PRO1083.

The entire nucleotide sequence of UNQ540 (DNA50921-1458) is shown in Figure 203 (SEQ ID NO:482). Clone UNQ540 (DNA50921-1458) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 214-216 and ending at the stop codon at nucleotide positions 2293-2295 (Figure 203). The

predicted polypeptide precursor is 693 amino acids long (Figure 204). The full-length PRO1083 protein shown in Figure 204 has an estimated molecular weight of about 77,738 daltons and a pI of about 8.87. Clone UNQ540 (DNA50921-1458) has been deposited with the ATCC. Regarding the sequence, it is understood that the deposited clone contains the correct sequence, and the sequences provided herein are based on known sequencing techniques.

Still analyzing the amino acid sequence of SEQ ID NO:483, the putative signal peptide is at about amino acids 1-25 of SEQ ID NO:483. The transmembrane domains are at about amino acids 382-398, 402-420, 445-468, 473-491, 519-537, 568-590 and 634-657 of SEQ ID NO:483. A microbodies C-terminal targeting signal is at about amino acids 691-693 of SEQ ID NO:483. cAMP- and cGMP-dependent protein kinase phosphorylation sites are at about amino acids 198-201 and 370-373 of SEQ ID NO:483. N-glycosylation sites are at about amino acids 39-42, 148-151, 171-174, 234-237, 303-306, 324-227 and 341-344 of SEQ ID NO:483. A G-protein coupled receptor family domain is at about amino acids 475-504 of SEQ ID NO:483. The corresponding nucleotides can be routinely determined given the sequences provided herein.

EXAMPLE 84: Isolation of cDNA Clones Encoding Human PRO200

Probes based on an expressed sequence tag (EST) identified from the Incyte Pharmaceuticals database due to homology with VEGF were used to screen a cDNA library derived from the human glioma cell line G61. In particular, Incyte Clone "INC1302516" was used to generate the following four probes:

(SEQ ID NO:489) ACTTCTCAGTGTCCATAAGGG;

(SEQ ID NO:490) GAACTAAAGAGAACCGATACCATTTTCTGGCCAGGTTGTC;

(SEQ ID NO:491) CACCACAGCGTTTAACCAGG; and

(SEQ ID NO:492) ACAACAGGCACAGTTCCCAC.

Nine positives were identified and characterized. Three clones contained the full coding region and were identical in sequence. Partial clones were also identified from a fetal lung library and were identical with the glioma-derived sequence with the exception of one nucleotide change which did not alter the encoded amino acid.

EXAMPLE 85: Expression Constructs for PRO200

For mammalian protein expression, the entire open reading frame (ORF) was cloned into a CMV-based expression vector. An epitope-tag (FLAG, Kodak) and Histidine-tag (His8) were inserted between the ORF and stop codon. VEGF-E-His8 and VEGF-E-FLAG were transfected into human embryonic kidney 293 cells by SuperFect (Qiagen) and pulse-labeled for 3 hours with [^{35}S]methionine and [^{3} C]cysteine. Both epitope-tagged proteins co-migrate when 20 microliters of 15-fold concentrated serum-free conditioned medium were electrophoresed on a polyacrylamide gel (Novex) in sodium dodecyl sulfate sample buffer (SDS-PAGE). The VEGF-E-IgG expression plasmid was constructed by cloning the ORF in front of the human Fc (IgG) sequence.

The VEGF-E-IgG plasmid was co-transfected with Baculogold Baculovirus DNA (Pharmingen) using Lipofectin (GibcoBRL) into 10^5 Sf9 cells grown in Hink's TNM-FH medium (JRH Biosciences) supplemented with 10% fetal bovine serum. Cells were incubated for 5 days at 28°C. The supernatant was harvested and subsequently used for the first viral amplification by infecting Sf9 cells at an approximate multiplicity of infection (MOI) of 10. Cells were incubated for 3 days, then supernatant harvested, and expression of the recombinant plasmid determined by binding of 1 ml of supernatant to 30 μl of Protein-A Sepharose CL-4B beads (Pharmacia) followed by subsequent

SDS-PAGE analysis. The first amplification supernatant was used to infect a 500 ml spinner culture of Sf9 cells grown in ESF-921 medium (Expression Systems LLC) at an approximate MOI of 0.1. Cells were treated as above, except harvested supernatant was sterile filtered. Specific protein was purified by binding to Protein-A Sepharose 4 Fast Flow (Pharmacia) column.

Example 86: Northern Blot Analyses for PRO200

Blots of human poly(A)+ RNA from multiple adult and fetal tissues and tumor cell lines were obtained from Clontech (Palo Alto, CA). Hybridization was carried out using ^{32}P-labeled probes containing the entire coding region and washed in 0.1 x SSC, 0.1% SDS at 63°C.

VEGF-E mRNA was detectable in fetal lung, kidney, brain, liver and adult heart, placenta, liver, skeletal muscle, kidney, and pancreas. VEGF-E mRNA was also found in A549 lung adenocarcinoma and HeLa cervical adenocarcinoma cell lines.

Example 87: In Situ Hybridization of Human Fetal Tissue Sections for PRO200

Formalin-fixed, paraffin-embedded human fetal brain, liver, lower limb, small intestine, thyroid, lymph node, thymus, stomach, trachea, skin, spleen, spinal cord, adrenal, placenta, cord, and adult liver, pancreas, lung, spleen, lymph node, adrenal, heart, aorta, and skin were sectioned, deparaffinized, deproteinated in proteinase K (20 μg/ml) for 15 minutes at 37°C, and further processed for in situ hybridization as described by Lu LH and Gillett NA (Cell Vision 1:169-176, 1994). A [α-^{33}-P]UTP-labeled antisense riboprobe was generated from a PCR product of 980 bp (primers GGCGGAATCCAACCTGAGTAG and GCGGCTATCCTCCTGTGCTC, SEQ ID NOS: 493 and 494, respectively). The slides were dipped in Kodak NTB2 nuclear track emulsion and exposed for 4 weeks.

VEGF-E mRNA expression included localization at the growth plate region and embracing fetal myocytes.

Example 88: Myocyte Hypertrophy Assay for PRO200

Myocytes from neonatal Harlan Sprague Dawley rat heart ventricle (23 days gestation) were plated in duplicate at 75000 cells/ml in a 96-well plate. Cells were treated for 48h with 2000, 200, 20, or 2 ng/ml VEGF-E-IgG. Myocytes were stained with crystal violet to visualize morphology and scored on a scale of 3 to 7, 3 being nonstimulated and 7 being full-blown hypertrophy.

2000 ng/ ml and 200 ng/ ml VEGF-E caused hypertrophy, scored as a 5.

Example 89: Cell Proliferation Assay for PRO200

Mouse embryonic fibroblast C3H1OT1/2 cells (ATCC) were grown in 50:50 Ham's F-12: low glucose DMEM medium containing 10% fetal calf serum (FCS). Cells were plated in duplicate in a 24-well plate at 1000, 2000, and 4000 cells/well. After 48 hours, cells were switched to medium containing 2% FCS and were incubated for 72 hours with 200, 800, or 2000 ng/ml VEGF-E or no growth factor added.

Approximately 1.5 fold greater number of cells were measured in the presence of 200 ng/ml VEGF-E as in its absence, at all three cell densities.

Example 90: Endothelial Cell Survival Assay for PRO200

Human umbilical vein endothelial cells (HUVEC, Cell Systems) were maintained in Complete Media (Cell Systems) and plated in triplicate in serum-free medium (Basic Media from Cell Systems containing 0.1% BSA) at 20,000 cells/well of a 48-well plate. Cells were incubated for 5 days with 200 or 400 ng/ml VEGF-E-IgG, 100 ng/ml VEGF, 20 ng/ml basic FGF, or no addition.

Survival was 2-3 times greater with VEGF-E as compared to lack of growth factor addition. VEGF and basic FGF were included as positive controls.

EXAMPLE 91: Isolation of cDNA Clones Encoding Human PRO285

A proprietary expressed sequence tag (EST) DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) was searched and an EST (#2243209) was identified which showed homology to the *Drosophila* Toll protein.

Based on the EST, a pair of PCR primers (forward and reverse):

TAAAGACCCAGCTGTGACCG (SEQ ID NO:499)

ATCCATGAGCCTCTGATGGG (SEQ ID NO: 500), and

a probe:

ATTTATGTCTCGAGGAAAGGGACTGGTTACCAGGGCAGCCAGTTC (SEQ ID NO: 501)

were synthesized.

mRNA for construction of the cDNA libraries was isolated from human placenta tissue. The cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA (Fast Track 2). The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into the cloning vector pCR2.1 (Invitrogen, Inc.) using reagents and protocols from Life Technologies, Gaithersburg, MD (Super Script Plasmid System). The double stranded cDNA was sized to greater than 1000 bp and the cDNA was cloned into BamHI/NotI cleaved vector. pCR2.1 is a commercially available plasmid, designed for easy cloning of PCR fragments, that carries AmpR and KanR genes for selection, and LacZ gene for blue-white selection.

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO285 gene using the probe oligonucleotide and one of the PCR primers.

A cDNA clone was sequenced in entirety. The entire nucleotide sequence of DNA40021-1154 (encoding PRO285) is shown in Figure 208 (SEQ ID NO:495). Clone DNA40021-1154 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 61-63 (Figure 208). The predicted polypeptide precursor is 1049 amino acids long, including a putative signal peptide at amino acid positions 1-29, a putative transmembrane domain between amino acid positions 837-860, and a leucine zipper pattern at amino acid positions 132-153 and 704-725, respectively. It is noted that the indicated boundaries are approximate, and the actual limits of the indicated regions might differ by a few amino acids. Clone DNA40021-1154 has been deposited with ATCC (designation: DNA40021-1154) and is assigned ATCC deposit no.209389.

Based on a BLAST and FastA sequence alignment analysis (using the ALIGN computer program) of the full-length sequence is a human analogue of the *Drosophila* Toll protein, and is homologous to the following human Toll proteins: Toll1 (DNAX# HSU88540-1, which is identical with the random sequenced full-length cDNA #HUMRSC786-1); Toll2 (DNAX# HSU88878-1); Toll3 (DNAX# HSU88879-1); and Toll4 (DNAX# HSU88880-1).

EXAMPLE 92: Isolation of cDNA Clones Encoding Human PRO286

A proprietary expressed sequence tag (EST) DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) was searched and an EST (#694401) was identified which showed homology to the *Drosophila* Toll protein.

Based on the EST, a pair of PCR primers (forward and reverse):

GCCGAGACAAAAACGTTCTCC (SEQ ID NO:502)

CATCCATGTTCTCATCCATTAGCC (SEQ ID NO: 503), and

a probe:

TCGACAACCTCATGCAGAGCATCAACCAAAGCAAGAAAACAGTATT (SEQ ID NO: 504)

were synthesized.

mRNA for construction of the cDNA libraries was isolated from human placenta tissue. This RNA was used to generate an oligo dT primed cDNA library in the vector pRK5D using reagents and protocols from Life Technologies, Gaithersburg, MD (Super Script Plasmid System). pRK5D is a cloning vector that has an sp6 transcription initiation site followed by an SfiI restriction enzyme site preceding the XhoI/NotI cDNA cloning sites. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized to greater than 1000 bp appropriately by gel electrophoresis, and cloned in a defined orientation into XhoI/NotI-cleaved pRK5D.

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO286 gene using the probe oligonucleotide identified above and one of the PCR primers.

A cDNA clone was sequenced in entirety. The entire nucleotide sequence of DNA42663-1154 (encoding PRO286) is shown in Figures 210A-B (SEQ ID NO:497). Clone DNA42663-1154 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 57-59 (Figure 211). The predicted polypeptide precursor is 1041 amino acids long, including a putative signal peptide at amino acid positions 1-26, a potential transmembrane domain at amino acid positions 826-848, and leucine zipper patterns at amino acids 130-151, 206-227, 662-684, 669-690 and 693-614, respectively. It is noted that the indicated boundaries are approximate, and the actual limits of the indicated regions might differ by a few amino acids. Clone DNA42663-1154 has been deposited with ATCC (designation: DNA42663-1154) and is assigned ATCC deposit no. 209386.

Based on a BLAST and FastA sequence alignment analysis (using the ALIGN computer program) of the full-length sequence of PR0286, it is a human analogue of the *Drosophila* Toll protein, and is homologous to the following human Toll proteins: Toll1 (DNAX# HSU88540-1, which is identical with the random sequenced full-length cDNA #HUMRSC786-1); Toll2 (DNAX# HSU88878-1); Toll3 (DNAX# HSU88879-1); and Toll4 (DNAX# HSU88880-1).

Example 93: NF-κB Assay for PRO285 and PRO286

As the Toll proteins signal through the NF-κB pathway, their biological activity can be tested in an NF-κB assay. In this assay Jurkat cells are transiently transfected using Lipofectamine reagent (Gibco BRL) according to the manufacturer's instructions. 1μg pB2XLuc plasmid, containing NF-κB-driven luciferase gene, is contransfected with 1μg pSRαN expression vector with or without the insert encoding PRO285 or PRO286. For a positive control, cells are treated with PMA (phorbol myristyl acetate; 20 ng/ml) and PHA (phytohaemaglutinin, 2μg/ml) for three to four hours. Cells are lysed 2 or 3 days later for measurement of luciferase activity using reagents from Promega.

EXAMPLE 94: Isolation of cDNA Clones Encoding Human PRO213-1, PRO1330 and PRO1449

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA28735. Based on the DNA28735 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO213-1, PRO1330 and/or PRO1449. A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-TGGAGCAGCAATATGCCAGCC-3' (SEQ ID NO:511)

reverse PCR primer 5'-TTTTCCACTCCTGTCGGGTTGG-3' (SEQ ID NO:512)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA28735 sequence which had the following nucleotide sequence:

hybridization probe

5'-GGTGACACTTGCCAGTCAGATGTGGATGAATGCAGTGCTAGGAGGG-3' (SEQ ID NO:513)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO213-1, PRO1330 and/or PRO1449 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal lung tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence encoding PRO213-1, PRO1330 and/or PRO1449 [DNA30943-1-1163-1 (SEQ ID NO:505), DNA64907-1163-1 (SEQ ID NO:507) and DNA64908-1163-1 (SEQ ID NO:509), respectively].

The entire nucleotide sequences corresponding to DNA30943-1-1163-1 (SEQ ID NO:505), DNA64907-1163-1 (SEQ ID NO:507) and DNA64908-1163-1 (SEQ ID NO:509), respectively. DNA30943-1163, DNA64907-1163-1 and DNA64908-1163-1 contain a single open reading frame with an apparent translational initiation site at nucleotide positions 336-338, 488-490 and 326-328, respectively, and ending at the stop codon at nucleotide positions 1221-1223, 1307-1309 and 1145-1147, respectively (Figures 212, 214 and 216). The predicted polypeptide precursor is 295, 273 and 273 amino acids long, respectively (Figures 213, 215 and 217). DNA30943-1-1163-1, DNA64907-1163-1 and DNA64908-1163-1 have been deposited with ATCC and are assigned ATCC deposit no. 209791, 203242 and 203243, respectively.

Analysis of the amino acid sequence of the full-length PRO213-1 polypeptide suggests that a portion of it possess significant homology to the human growth arrest-specific gene 6 protein. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced significant homology between the PRO213 amino acid sequence and the following Dayhoff sequences, HSMHC3W5A_6 and B48089.

Additional analysis of the amino acid sequence of the full-length PRO1330 and PRO1449 polypeptide indicates significant identity with notch4. More specifically, an analysis of the Dayhoff database (version 35.130 SwissProt 35) evidenced significant identity between PRO1330 and the following Dayhoff sequences, D86566_1 and NEL_HUMAN.

EXAMPLE 95: Isolation of cDNA Clones Encoding Human PRO298

A cDNA isolated in the amylase screen described in Example 2 above is herein designated DNA26832 (Figure 220; SEQ ID NO:516). The sequence of DNA26832 was then used to search expressed sequence tag (EST) databases. The EST databases included public EST databases (e.g., GenBank) and a proprietary EST database

(LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA). The search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266: 469-480 [1996]). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode proteins were clustered and assembled into consensus DNA sequences with the program "phrap" (Phil Green, University of Washington, Seattle, Washington; http://bozeman.mbt.washington.edu/phrap.docs/phrap.html).

A consensus DNA sequence was assembled relative to other EST sequences using phrap. A consensus sequence was determined, which was then extended using repeated cycles of BLAST and phrap to extend the consensus sequence as far as possible using the sources of EST sequences discussed above. The extended assembly sequence was designated DNA35861.Based on the DNA35861 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence of PRO298. Forward and reverse primers generally range from 20 to 30 nucleotides and are often designed to give a PCR product of about 100-1000 bp in length. The probe sequence is typically 40-55 bp in length. In some cases, additional oligonucleotides are synthesized when the consensus sequence is greater than about 1-1.5kbp. In order to screen several libraries for a full-length clone, DNA from the libraries was screened by PCR amplification, as per Ausubel et al., Current Protocols in Molecular Biology, with the PCR primer pair. A positive library was used to isolate clones encoding the gene of interest using the probe oligonucleotide and one of the primer pairs.

PCR primers (forward and reverse) and a hybridization probe were synthesized:
forward PCR primer 1 CAACGTGATTTCAAAGCTGGGCTC (SEQ ID NO:517)
forward PCR primer 2 GCCTCGTATCAAGAATTTCC (SEQ ID NO:518)
forward PCR primer 3 AGTGGAAGTCGACCTCCC (SEQ ID NO:519)
reverse PCR primer 1 CTCACCTGAAATCTCTCATAGCCC (SEQ ID NO:520)
hybridization probe 1 CGCAAAACCCATTTTGGGAGCAGGAATTCCAATCATGTCTGTGATGGTGG (SEQ ID NO:521)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO298 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue (LIB25). The cDNA libraries used to isolated the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO298 (herein designated UNQ261 [DNA39975-1210]) (SEQ ID NO:514), and the derived protein sequence for PRO298 (SEQ ID NO:515).

The entire nucleotide sequence of UNQ261 (DNA39975-1210) is shown in Figure 218 (SEQ ID NO:514). Clone DNA39975-1210 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 375-377. The predicted polypeptide precursor is 364 amino acids long. The protein contains four putative

transmembrane domains between amino acid positions 36-55 (type II TM), 65-84, 188-208, and 229-245, respectively. A putative N-linked glycosylation site starts at amino acid position 253. In addition, the following features have been identified in the protein sequence: cAMP- and cGMP-dependent protein kinase phosphorylation site, starting at position 8; N-myristoylation sites starting a position 173 and 262, respectively; and a ZP domain between amino acid positions 45-60. Clone DNA39975-1210 has been deposited with ATCC (April 21, 1998) and is assigned ATCC deposit no.209783.

EXAMPLE 96: Isolation of cDNA Clones Encoding Human PRO337

A cDNA sequence identified in the amylase screen described in Example 2 above is herein designated DNA42301 (Figure 223, SEQ ID NO:524). The DNA42301 sequence was then compared to other EST sequences using phrap as described in Example 1 above and a consensus sequence designated herein as DNA28761 was identified. Based on this consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence. In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO337 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal brain.

A cDNA clone was sequenced in its entirety. The full length nucleotide sequence of DNA43316-1237 is shown in Figure 221 (SEQ ID NO:522). Clone DNA43316-1237 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 134-136 (Figure 221; SEQ ID NO:522). The predicted polypeptide precursor is 344 amino acids long. Clone DNA43316-1237 has been deposited with ATCC and is assigned ATCC deposit no. 209487

Based on a BLAST-2 and FastA sequence alignment analysis of the full-length sequence, PRO337 shows amino acid sequence identity to rat neurotrimin (97%).

EXAMPLE 97: Isolation of cDNA Clones Encoding Human PRO403

Introduction:

Human thrombopoietin (THPO) is a glycosylated hormone of 352 amino acids consisting of two domains. The N-terminal domain, sharing 50% similarity to erythropoietin, is responsible for the biological activity. The C-terminal region is required for secretion. The gene for thrombopoietin (THPO) maps to human chromosome 3q27-q28 where the six exons of this gene span 7 kilobase base pairs of genomic DNA (Chang et al., Genomics 26: 636-7 (1995); Foster et al., Proc. Natl. Acad. Sci. USA 91: 13023-7 (1994); Gurney et al., Blood 85: 981-988 (1995). In order to determine whether there were any genes encoding THPO homologues located in close proximity to THPO, genomic DNA fragments from this region were identified and sequenced. Three P1 clones and one PAC clones (Genome Systems Inc., St. Louis, MO; cat. Nos. P1-2535 and PAC-6539) encompassing the THPO locus were isolated and a 140 kb region was sequenced using the ordered shotgun strategy (Chen et al., Genomics 17: 651-656 (1993)), coupled with a PCR-based gap filling approach. Analysis reveals that the region is gene-rich with four additional genes located very close to THPO: tumor necrosis factor-receptor type 1 associated protein 2 (TRAP2) and elongation initiation factor gamma (eIF4G), chloride channel 2 (CLCN2) and RNA polymerase II subunit hRPB17. While no THPO homolog was found in the region, four novel genes have been predicted by computer-assisted gene

detection (GRAIL)(Xu et al., Gen. Engin. 16: 241-253 (1994), the presence of CpG islands (Cross, S. and Bird, A., Curr. Opin. Genet. & Devel. 5: 109-314 (1995), and homology to known genes (as detected by WU-BLAST2.0)(Altschul and Gish, Methods Enzymol. 266: 460-480 (1996) (http://blast.wustl.edu/blast/README.html).

Procedures:

P1 and PAC clones:

The initial human P1 clone was isolated from a genomic P1 library (Genome Systems Inc., St. Louis, MO; cat. no.: P1-2535) screened with PCR primers designed from the THPO genomic sequence (A.L. Gurney, et al., Blood 85: 981-88 (1995). PCR primers were designed from the end sequences derived from this P1 clone were then used to screen P1 and PAC libraries (Genome Systems, Cat. Nos.: P1-2535 & PAC-6539) to identify overlapping clones (PAC1, p1.t, and P1.u). The 3'-end sequence from PAC.z was used to define the primers used for the screening of a human BAC library (Genome Systems Inc., St. Louis, MO; Cat. No.: BDTW-4533A).

Ordered Shotgun Strategy:

The Ordered Shotgun Strategy (OSS) (Chen et al., Genomics 17: 651-656 (1993)) involves the mapping and sequencing of large genomic DNA clones with a hierarchical approach. The P1 or PAC clone was sonicated and the fragments subcloned into lambda vector (λBluestar) (Novagen, Inc., Madison, WI: cat. no. 69242-3). The lambda subclone inserts were isolated by long-range PCR (Barnes, W. Proc. Natl. Acad. Sci. USA 91: 2216-2220 (1994) and the ends sequenced. The lambda-end sequences were overlapped to create a partial map of the original clone. Those lambda clones with overlapping end-sequences were identified, the insets subcloned into a plasmid vector (pUC18 or pUC19, Hoefer Pharmacia Biotech, Inc., San Francisco, CA, Cat. Nos. 27-4949-01 and 27-4951-01) and the ends of the plasmid subclones were sequenced and assembled to generate a contiguous sequence. This directed sequencing strategy minimizes the redundancy required while allowing one to scan for and concentrate on interesting regions.

In order to define better the THPO locus and to search for other genes related to the hematopoietin family, five genomic clones were isolated from this region by PCR screening of human P1 and PAC libraries (Genome System, Inc., Cat. Nos.: P1-2535 and PAC-6539).

The sizes of the genomic fragments are as follows: P1.t is 40 kb; P1.g is 70 kb; P1.u is 70 kb; PAC.z is 200 kb; and BAC.1 is 80 kb. Approximately 75% (140 kb) of the 190 kb genomic DNA region was sequenced by the Ordered Shotgun Strategy (OSS) (Chen et al., Genomics 17: 651-56 (1993), and assembled into contigs using AutoAssemblerTM (Applied Biosystems, Perkin Elmer, Foster City, CA, cat. no. 903227). The preliminary order of these contigs was determined by manual analysis. There were 47 contigs the 140 kb region. A PCR-based approach to ordering the contigs and filling in the gaps was employed. The following summarizes the number and sizes of the gaps. The 50 kb of sequence unique to BAC.1 was sequenced by a total shotgun approach with a ten-fold redundancy.

Size of gap	number
<50 bp	13
50-150 bp	7
150-300 bp	7
300-1000 bp	10
1000-5000 bp	7
> 5000 bp	2 ((15,000 bp)

DNA sequencing:

ABI DYE-primerTM chemistry (PE Applied Biosystems, Foster City, CA; Cat. No.: 402112) was used to end-sequence the lambda and plasmid subclones. ABI DYE-terminaterTM chemistry (PE Applied Biosystems, Foster City, CA, Cat. No: 403044) was used to sequence the PCR products with their respective PCR primers. The sequences were collected with an ABI377 instrument. For PCR products larger than 1kb, walking primers were used. The sequences of contigs generated by the OSS strategy in AutoAssemblerTM (PE Applied Biosystems, Foster City, CA; Cat. No: 903227) and the gap-filling sequencing trace files were imported into SequencherTM (Gene Codes Corp., Ann Arbor, MI) for overlapping and editing. The sequences generated by the total shotgun strategy were assembled using Phred and Phrap and edited using Consed (http://chimera.biotech.washington.edu/uwgc/projects.htm) and GFP (Genome Reconstruction Manager for Phrap), version 1.2 (http://stork.cellb.bcm.tmc.edu/gfp/).

PCR-Based gap filling Strategy:

Primers were designed based on the 5'- and 3'-end sequenced of each contig, avoiding repetitive and low quality sequence regions. All primers were designed to be 19-24-mers with 50-70% G/C content. Oligos were synthesized and gel-purified by standard methods.

Since the orientation and order of the contigs were unknown, permutations of the primers were used in the amplification reactions. Two PCR kits were used: first, XL PCR kit (Perkin Elmer, Norwalk, CT; Cat. No.: N8080205), with extension times of approximately 10 minutes; and second, the Taq polymerase PCR kit (Qiagen Inc., Valencia, CA; Cat. No.: 201223) was used under high stringency conditions if smeared or multiple products were observed with the XL PCR kit. The main PCR product from each successful reaction was extracted from a 0.9% low melting agarose gel and purified with the Geneclean DNA Purification kit prior to sequencing.

Analysis:

The identification and characterization of coding regions was carried out as follows: First, repetitive sequences were masked using RepeatMasker (A.F.A. Smit & P. Green, http://ftp.genome.washington.edu/RM/RM_details.html) which screens DNA sequences in FastA format against a library of repetitive elements and returns a masked query sequence. Repeats not masked were identified by comparing the sequence to the GenBank database using WUBLAST2.0 [Altschul, S & Gish, W., Methods Enzymol. 266: 460-480 (1996); http://blast.wustl.edu/blast/README.html] and were masked manually.

Next, known genes were revealed by comparing the genomic regions against Genentech's protein database using the WUBLAST2.0 algorithm and then annotated by aligning the genomic and cDNA sequences for each gene, respectively, using a Needleman-Wunch (Needleman and Wunsch, J. Mol. Biol. 48: 443-453 (1970) algorithm to find regions of local identity between sequences. The strategy results in detection of all exons of the five known genes in the region, THPO, TRAP2, eIF4g, CLCN2 and hRPB17 (see below).

Known genes	Map position
eukaryotic translation initiation factor 4 gamma	3q27-qter
thrombopoietin	3q26-q27
chloride channel 2	3q26-qter
TNF receptor associated protein 2	not previously mapped
RNA polymerase II subunit hRPB17	not previously mapped

Finally, novel transcription units were predicted using a number of approaches. CpG islands (S. Cross & Bird, A., Curr. Opin. Genet. Dev. 5: 109-314 (1995) islands were used to define promoter regions and were identified as clusters of sites cleaved by enzymes recognizing GC-rich, 6 or 8-mer palindromic sequences (NotI, NarI, BssHII, XhoI. CpG islands are usually associated with promoter regions of genes. WUBLAST2.0 analysis of short genomic regions (10-20 kb) versus GenBank revealed matches to ESTs. The individual EST sequences (or where possible, their sequence chromatogram files) were retrieved and assembled with Sequencer to provide a theoretical cDNA sequence (DNA36443). GRAIL2 (ApoCom Inc., Knoxville, TN, command line version for the DEC alpha) was used to predict a novel exon. The five known genes in the region served as internal controls for the success of the GRAIL algorithm.

Isolation:

A partial endothelin converting enzyme-2 (ECE-2) cDNA clone was isolated by first splicing in silico the ECE-2 exons predicted in the genomic sequence to generate a putative sequence (DNA36443). An oligonucleotide probe: GAAGCAGTGCAGCCAGCAGTAGAGAGGCACCTGCTAAGA) (SEQ ID NO:530) was designed and used to screen a human fetal small intestine library (LIB110) and internal PCR primers (36443f1) (ECE2.f:ACGCAGCTGGAGCTGGTCTTAGCA) (SEQ ID NO:531) and (36443r1) (ECE2.r) (GGTACTGGACCCCTAGGGCCACAA) (SEQ ID NO:532) were used to confirm clones hybridizing to the probe prior to sequencing. One positive clone was obtained, however this cDNA (DNA49830) represented a partially spliced transcript containing appropriately spliced exons 1 through 6, followed by intron 6 sequence. The oligo dT primer annealed to a polyA-stretch within an Alu element present in intron 6. An additional ECE-2 cDNA fragment (DNA49831) was obtained by PCR from a human fetal kidney library (LIB227) with primers designed from the presumed cDNA sequence [36443f3: CCTCCCAGCCGAGACCAGTGG (SEQ ID NO:533) and 36443r2: GGTCCTATAAGGGCCAAGACC (SEQ ID NO:534)]. This PCR product extended from exon 13 into the 3' untranslated region in exon 18.

A full length endothelin converting enzyme 2 (ECE-2) cDNA clone (DNA55800-1263) was isolated from an oligo-dT-primed human fetal brain library. RNA from human fetal brain tissue (20 weeks gestation, #283005)(SRC175) was isolated by guanidine thiocyanate and 5 μg used to generate double stranded cDNA which was cloned into the vector pRK5E. The 3' -primer (pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT) (SEQ ID NO:535) and the 5 -linker (pCGGACGCGTGGGTCGA) (SEQ ID NO:536) were designed to introduce XhoI and NotI restriction sites. The library was screened with PCR primers [36443pcrf1: CGGCCGTGATGGCTGGTGACG (SEQ ID NO:537) and 36443r3: GGCAGACTCCTTCCTATGGG (SEQ ID NO:538)] designed from the partial human ECE-2 cDNA sequences (DNA49830 and DNA49831). PCR products were cloned into the vector pCR2.1-TOPO (Invitrogen Corp., Carlsbad, CA, Cat. No. K4500-01) and sequenced with DYE-terminator chemistry as described above.

EXAMPLE 98: Northern Blot and in situ RNA Hybridization Analysis for PRO403

Expression of PRO403 mRNA in human tissues was examined by Northern blot analysis. Human polyA+ RNA blots derived from human fetal and adult tissues (Clontech, Palo Alto, CA; Cat. Nos. 7760-1, 7756-1 and 7755-1) were hybridized to a [32P-α]dATP-labelled cDNA fragments from probe based on the full length PRO403 cDNA. Blots were incubated with the probes in hybridization buffer (5X SSPE: 2X Denhardt's solution; 100 mg/mL denatured sheared salmon sperm DNA; 50% formamide; 2% SDS) for 18 hours at 42°C, washed to high stringency

(0.1XSSC, 0.1% SDS, 50°C) and autoradiographed. The blots were developed after overnight exposure by phosphorimager analysis (Fuji).

PRO403 mRNA transcripts were detected. Analysis of the expression pattern showed the strongest signal of the expected 3.3 kb transcript in adult brain (highest in the cerebellum, putamen, medulla, and temporal lobe, and lower in the cerebral cortex, occipital lobe and frontal lobe), spinal cord, lung and pancreas and higher levels of a 4.5 kb transcript in fetal brain and kidney.

EXAMPLE 99: Use of PRO Polypeptide-Encoding Nucleic Acid as Hybridization Probes

The following method describes use of a nucleotide sequence encoding a PRO polypeptide as a hybridization probe.

DNA comprising the coding sequence of of a PRO polypeptide of interest as disclosed herein may be employed as a probe or used as a basis from which to prepare probes to screen for homologous DNAs (such as those encoding naturally-occurring variants of the PRO polypeptide) in human tissue cDNA libraries or human tissue genomic libraries.

Hybridization and washing of filters containing either library DNAs is performed under the following high stringency conditions. Hybridization of radiolabeled PRO polypeptide-encoding nucleic acid-derived probe to the filters is performed in a solution of 50% formamide, 5x SSC, 0.1% SDS, 0.1% sodium pyrophosphate, 50 mM sodium phosphate, pH 6.8, 2x Denhardt's solution, and 10% dextran sulfate at 42°C for 20 hours. Washing of the filters is performed in an aqueous solution of 0.1x SSC and 0.1% SDS at 42°C.

DNAs having a desired sequence identity with the DNA encoding full-length native sequence PRO polypeptide can then be identified using standard techniques known in the art.

EXAMPLE 100: Expression of PRO Polypeptides in *E. coli*

This example illustrates preparation of an unglycosylated form of a desired PRO polypeptide by recombinant expression in *E. coli*.

The DNA sequence encoding the desired PRO polypeptide is initially amplified using selected PCR primers. The primers should contain restriction enzyme sites which correspond to the restriction enzyme sites on the selected expression vector. A variety of expression vectors may be employed. An example of a suitable vector is pBR322 (derived from *E. coli*; see Bolivar et al., Gene, 2:95 (1977)) which contains genes for ampicillin and tetracycline resistance. The vector is digested with restriction enzyme and dephosphorylated. The PCR amplified sequences are then ligated into the vector. The vector will preferably include sequences which encode for an antibiotic resistance gene, a trp promoter, a polyhis leader (including the first six STII codons, polyhis sequence, and enterokinase cleavage site), the specific PRO polypeptide coding region, lambda transcriptional terminator, and an argU gene.

The ligation mixture is then used to transform a selected *E. coli* strain using the methods described in Sambrook et al., supra. Transformants are identified by their ability to grow on LB plates and antibiotic resistant colonies are then selected. Plasmid DNA can be isolated and confirmed by restriction analysis and DNA sequencing.

Selected clones can be grown overnight in liquid culture medium such as LB broth supplemented with antibiotics. The overnight culture may subsequently be used to inoculate a larger scale culture. The cells are then grown to a desired optical density, during which the expression promoter is turned on.

After culturing the cells for several more hours, the cells can be harvested by centrifugation. The cell pellet obtained by the centrifugation can be solubilized using various agents known in the art, and the solubilized PRO polypeptide can then be purified using a metal chelating column under conditions that allow tight binding of the protein.

PRO181, PRO195, PRO200, PRO237, PRO273, PRO540, PRO322, PRO1017, PRO938, PRO162, PRO1114, PRO827 and PRO1008 were expressed in *E. coli* in a poly-His tagged form, using the following procedure. The DNA encoding the PRO polypeptide was initially amplified using selected PCR primers. The primers contained restriction enzyme sites which correspond to the restriction enzyme sites on the selected expression vector, and other useful sequences providing for efficient and reliable translation initiation, rapid purification on a metal chelation column, and proteolytic removal with enterokinase. The PCR-amplified, poly-His tagged sequences were then ligated into an expression vector, which was used to transform an *E. coli* host based on strain 52 (W3110 fuhA(tonA) lon galE rpoHts(htpRts) clpP(lacIq). Transformants were first grown in LB containing 50 mg/ml carbenicillin at 30°C with shaking until an O.D.600 of 3-5 was reached. Cultures were then diluted 50-100 fold into CRAP media (prepared by mixing 3.57 g $(NH_4)_2SO_4$, 0.71 g sodium citrate·2H2O, 1.07 g KCl, 5.36 g Difco yeast extract, 5.36 g Sheffield hycase SF in 500 mL water, as well as 110 mM MPOS, pH 7.3, 0.55% (w/v) glucose and 7 mM $MgSO_4$) and grown for approximately 20-30 hours at 30°C with shaking. Samples were removed to verify expression by SDS-PAGE analysis, and the bulk culture is centrifuged to pellet the cells. Cell pellets were frozen until purification and refolding.

E. coli paste from 0.5 to 1 L fermentations (6-10 g pellets) was resuspended in 10 volumes (w/v) in 7 M guanidine, 20 mM Tris, pH 8 buffer. Solid sodium sulfite and sodium tetrathionate is added to make final concentrations of 0.1M and 0.02 M, respectively, and the solution was stirred overnight at 4°C. This step results in a denatured protein with all cysteine residues blocked by sulfitolization. The solution was centrifuged at 40,000 rpm in a Beckman Ultracentifuge for 30 min. The supernatant was diluted with 3-5 volumes of metal chelate column buffer (6 M guanidine, 20 mM Tris, pH 7.4) and filtered through 0.22 micron filters to clarify. Depending the clarified extract was loaded onto a 5 ml Qiagen Ni-NTA metal chelate column equilibrated in the metal chelate column buffer. The column was washed with additional buffer containing 50 mM imidazole (Calbiochem, Utrol grade), pH 7.4. The protein was eluted with buffer containing 250 mM imidazole. Fractions containing the desired protein were pooled and stored at 4°C. Protein concentration was estimated by its absorbance at 280 nm using the calculated extinction coefficient based on its amino acid sequence.

The proteins were refolded by diluting sample slowly into freshly prepared refolding buffer consisting of: 20 mM Tris, pH 8.6, 0.3 M NaCl, 2.5 M urea, 5 mM cysteine, 20 mM glycine and 1 mM EDTA. Refolding volumes were chosen so that the final protein concentration was between 50 to 100 micrograms/ml. The refolding solution was stirred gently at 4°C for 12-36 hours. The refolding reaction was quenched by the addition of TFA to a final concentration of 0.4% (pH of approximately 3). Before further purification of the protein, the solution was filtered through a 0.22 micron filter and acetonitrile was added to 2-10% final concentration. The refolded protein was chromatographed on a Poros R1/H reversed phase column using a mobile buffer of 0.1% TFA with elution with a gradient of acetonitrile from 10 to 80%. Aliquots of fractions with A280 absorbance were analyzed on SDS polyacrylamide gels and fractions containing homogeneous refolded protein were pooled. Generally, the properly refolded species of most proteins are eluted at the lowest concentrations of acetonitrile since those species are the most compact with their hydrophobic interiors shielded from interaction with the reversed phase resin. Aggregated

species are usually eluted at higher acetonitrile concentrations. In addition to resolving misfolded forms of proteins from the desired form, the reversed phase step also removes endotoxin from the samples.

Fractions containing the desired folded PRO proteins were pooled and the acetonitrile removed using a gentle stream of nitrogen directed at the solution. Proteins were formulated into 20 mM Hepes, pH 6.8 with 0.14 M sodium chloride and 4% mannitol by dialysis or by gel filtration using G25 Superfine (Pharmacia) resins equilibrated in the formulation buffer and sterile filtered.

EXAMPLE 101: Expression of PRO Polypeptides in Mammalian Cells

This example illustrates preparation of a glycosylated form of a desired PRO polypeptide by recombinant expression in mammalian cells.

The vector, pRK5 (see EP 307,247, published March 15, 1989), is employed as the expression vector. Optionally, the PRO polypeptide-encoding DNA is ligated into pRK5 with selected restriction enzymes to allow insertion of the PRO polypeptide DNA using ligation methods such as described in Sambrook et al., supra. The resulting vector is called pRK5-PRO polypeptide.

In one embodiment, the selected host cells may be 293 cells. Human 293 cells (ATCC CCL 1573) are grown to confluence in tissue culture plates in medium such as DMEM supplemented with fetal calf serum and optionally, nutrient components and/or antibiotics. About 10 μg pRK5-PRO polypeptide DNA is mixed with about 1 μg DNA encoding the VA RNA gene [Thimmappaya et al., Cell, 31:543 (1982)] and dissolved in 500 μl of 1 mM Tris-HCl, 0.1 mM EDTA, 0.227 M $CaCl_2$. To this mixture is added, dropwise, 500 μl of 50 mM HEPES (pH 7.35), 280 mM NaCl, 1.5 mM $NaPO_4$, and a precipitate is allowed to form for 10 minutes at 25°C. The precipitate is suspended and added to the 293 cells and allowed to settle for about four hours at 37°C. The culture medium is aspirated off and 2 ml of 20% glycerol in PBS is added for 30 seconds. The 293 cells are then washed with serum free medium, fresh medium is added and the cells are incubated for about 5 days.

Approximately 24 hours after the transfections, the culture medium is removed and replaced with culture medium (alone) or culture medium containing 200 μCi/ml ^{35}S-cysteine and 200 μCi/ml ^{35}S-methionine. After a 12 hour incubation, the conditioned medium is collected, concentrated on a spin filter, and loaded onto a 15% SDS gel. The processed gel may be dried and exposed to film for a selected period of time to reveal the presence of PRO polypeptide. The cultures containing transfected cells may undergo further incubation (in serum free medium) and the medium is tested in selected bioassays.

In an alternative technique, PRO polypeptide may be introduced into 293 cells transiently using the dextran sulfate method described by Somparyrac et al., Proc. Natl. Acad. Sci., 12:7575 (1981). 293 cells are grown to maximal density in a spinner flask and 700 μg pRK5-PRO polypeptide DNA is added. The cells are first concentrated from the spinner flask by centrifugation and washed with PBS. The DNA-dextran precipitate is incubated on the cell pellet for four hours. The cells are treated with 20% glycerol for 90 seconds, washed with tissue culture medium, and re-introduced into the spinner flask containing tissue culture medium, 5 μg/ml bovine insulin and 0.1 μg/ml bovine transferrin. After about four days, the conditioned media is centrifuged and filtered to remove cells and debris. The sample containing expressed PRO polypeptide can then be concentrated and purified by any selected method, such as dialysis and/or column chromatography.

In another embodiment, PRO polypeptides can be expressed in CHO cells. The pRK5-PRO polypeptide can be transfected into CHO cells using known reagents such as $CaPO_4$ or DEAE-dextran. As described above, the

cell cultures can be incubated, and the medium replaced with culture medium (alone) or medium containing a radiolabel such as ^{35}S-methionine. After determining the presence of PRO polypeptide, the culture medium may be replaced with serum free medium. Preferably, the cultures are incubated for about 6 days, and then the conditioned medium is harvested. The medium containing the expressed PRO polypeptide can then be concentrated and purified by any selected method.

Epitope-tagged PRO polypeptide may also be expressed in host CHO cells. The PRO polypeptide may be subcloned out of the pRK5 vector. The subclone insert can undergo PCR to fuse in frame with a selected epitope tag such as a poly-his tag into a Baculovirus expression vector. The poly-his tagged PRO polypeptide insert can then be subcloned into a SV40 driven vector containing a selection marker such as DHFR for selection of stable clones. Finally, the CHO cells can be transfected (as described above) with the SV40 driven vector. Labeling may be performed, as described above, to verify expression. The culture medium containing the expressed poly-His tagged PRO polypeptide can then be concentrated and purified by any selected method, such as by Ni^{2+}-chelate affinity chromatography.

Stable expression in CHO cells was performed using the following procedure. The proteins were expressed as an IgG construct (immunoadhesin), in which the coding sequences for the soluble forms (e.g. extracellular domains) of the respective proteins were fused to an IgG1 constant region sequence containing the hinge, CH2 and CH2 domains and/or is a poly-His tagged form.

Following PCR amplification, the respective DNAs were subcloned in a CHO expression vector using standard techniques as described in Ausubel et al., *Current Protocols of Molecular Biology*, Unit 3.16, John Wiley and Sons (1997). CHO expression vectors are constructed to have compatible restriction sites 5' and 3' of the DNA of interest to allow the convenient shuttling of cDNA's. The vector used expression in CHO cells is as described in Lucas *et al.*, *Nucl. Acids Res.* 24: 9 (1774-1779 (1996), and uses the SV40 early promoter/enhancer to drive expression of the cDNA of interest and dihydrofolate reductase (DHFR). DHFR expression permits selection for stable maintenance of the plasmid following transfection.

Twelve micrograms of the desired plasmid DNA were introduced into approximately 10 million CHO cells using commercially available transfection reagents Superfect® (Quiagen), Dosper® or Fugene® (Boehringer Mannheim). The cells were grown and described in Lucas *et al.*, supra. Approximately 3 x 10^{-7} cells are frozen in an ampule for further growth and production as described below.

The ampules containing the plasmid DNA were thawed by placement into water bath and mixed by vortexing. The contents were pipetted into a centrifuge tube containing 10 mLs of media and centrifuged at 1000 rpm for 5 minutes. The supernatant was aspirated and the cells were resuspended in 10 mL of selective media (0.2 μm filtered PS20 with 5% 0.2 μm diafiltered fetal bovine serum). The cells were then aliquoted into a 100 mL spinner containing 90 mL of selective media. After 1-2 days, the cells were transferred into a 250 mL spinner filled with 150 mL selective growth medium and incubated at 37°C. After another 2-3 days, a 250 mL, 500 mL and 2000 mL spinners were seeded with 3 x 10^5 cells/mL. The cell media was exchanged with fresh media by centrifugation and resuspension in production medium. Although any suitable CHO media may be employed, a production medium described in US Patent No. 5,122,469, issued June 16, 1992 was actually used. 3L production spinner is seeded at 1.2 x 10^6 cells/mL. On day 0, the cell number pH were determined. On day 1, the spinner was sampled and sparging with filtered air was commenced. On day 2, the spinner was sampled, the temperature shifted to 33°C, and 30 mL of 500 g/L glucose and 0.6 mL of 10% antifoam (e.g., 35% polydimethylsiloxane emulsion, Dow Corning

365 Medical Grade Emulsion). Throughout the production, pH was adjusted as necessary to keep at around 7.2. After 10 days, or until viability dropped below 70%, the cell culture was harvested by centrifugtion and filtering through a 0.22 μm filter. The filtrate was either stored at 4°C or immediately loaded onto columns for purification.

For the poly-His tagged constructs, the proteins were purified using a Ni-NTA column (Qiagen). Before purification, imidazole was added to the conditioned media to a concentration of 5 mM. The conditioned media was pumped onto a 6 ml Ni-NTA column equilibrated in 20 mM Hepes, pH 7.4, buffer containing 0.3 M NaCl and 5 mM imidazole at a flow rate of 4-5 ml/min. at 4°C. After loading, the column was washed with additional equilibration buffer and the protein eluted with equilibration buffer containing 0.25 M imidazole. The highly purified protein was subsequently desalted into a storage buffer containing 10 mM Hepes, 0.14 M NaCl and 4% mannitol, pH 6.8, with a 25 ml G25 Superfine (Pharmacia) column and stored at -80°C.

Immunoadhesin (Fc containing) constructs of were purified from the conditioned media as follows. The conditioned medium was pumped onto a 5 ml Protein A column (Pharmacia) which had been equilibrated in 20 mM Na phosphate buffer, pH 6.8. After loading, the column was washed extensively with equilibration buffer before elution with 100 mM citric acid, pH 3.5. The eluted protein was immediately neutralized by collecting 1 ml fractions into tubes containing 275 μL of 1 M Tris buffer, pH 9. The highly purified protein was subsequently desalted into storage buffer as described above for the poly-His tagged proteins. The homogeneity was assessed by SDS polyacrylamide gels and by N-terminal amino acid sequencing by Edman degradation.

The following PRO polypeptides were successfully transiently expressed in CHO cells: PRO200, PRO320, PRO237, PRO273, PRO337, PRO846, PRO363, PRO322, PRO1083, PRO938, PRO1012, PRO1114, PRO1008 and PRO1075.

The following PRO polypeptides were successfully transiently expressed in COS cells: PRO181, PRO195, PRO200, PRO320, PRO237, PRO273, PRO285, PRO337, PRO526, PRO540, PRO846, PRO362, PRO363, PRO700, PRO707, PRO617, PRO322, PRO719, PRO1083, PRO868, PRO866, PRO768, PRO938, PRO1012, PRO162, PRO1114, PRO827, PRO1008 and PRO1075.

The following PRO polypeptides were successfully stably expressed in CHO cells: PRO181, PRO195, PRO200, PRO320, PRO285, PRO337, PRO846, PRO362, PRO363, PRO707, PRO617, PRO322, PRO1083, PRO868, PRO866, PRO1017, PRO792, PRO788, PRO938, PRO1012, PRO162, PRO1114, PRO827, PRO1008, PRO1075 and PRO1031.

EXAMPLE 102: Expression of PRO Polypeptides in Yeast

The following method describes recombinant expression of a desired PRO polypeptide in yeast.

First, yeast expression vectors are constructed for intracellular production or secretion of PRO polypeptides from the ADH2/GAPDH promoter. DNA encoding a desired PRO polypeptide, a selected signal peptide and the promoter is inserted into suitable restriction enzyme sites in the selected plasmid to direct intracellular expression of the PRO polypeptide. For secretion, DNA encoding the PRO polypeptide can be cloned into the selected plasmid, together with DNA encoding the ADH2/GAPDH promoter, the yeast alpha-factor secretory signal/leader sequence, and linker sequences (if needed) for expression of the PRO polypeptide.

Yeast cells, such as yeast strain AB110, can then be transformed with the expression plasmids described above and cultured in selected fermentation media. The transformed yeast supernatants can be analyzed by precipitation with 10% trichloroacetic acid and separation by SDS-PAGE, followed by staining of the gels with

Coomassie Blue stain.

Recombinant PRO polypeptide can subsequently be isolated and purified by removing the yeast cells from the fermentation medium by centrifugation and then concentrating the medium using selected cartridge filters. The concentrate containing the PRO polypeptide may further be purified using selected column chromatography resins.

EXAMPLE 103: Expression of PRO Polypeptides in Baculovirus-Infected Insect Cells

The following method describes recombinant expression of PRO polypeptides in Baculovirus-infected insect cells.

The desired PRO polypeptide is fused upstream of an epitope tag contained with a baculovirus expression vector. Such epitope tags include poly-his tags and immunoglobulin tags (like Fc regions of IgG). A variety of plasmids may be employed, including plasmids derived from commercially available plasmids such as pVL1393 (Novagen). Briefly, the PRO polypeptide or the desired portion of the PRO polypeptide (such as the sequence encoding the extracellular domain of a transmembrane protein) is amplified by PCR with primers complementary to the 5' and 3' regions. The 5' primer may incorporate flanking (selected) restriction enzyme sites. The product is then digested with those selected restriction enzymes and subcloned into the expression vector.

Recombinant baculovirus is generated by co-transfecting the above plasmid and BaculoGold™ virus DNA (Pharmingen) into *Spodoptera frugiperda* ("Sf9") cells (ATCC CRL 1711) using lipofectin (commercially available from GIBCO-BRL). After 4-5 days of incubation at 28°C, the released viruses are harvested and used for further amplifications. Viral infection and protein expression is performed as described by O'Reilley et al., Baculovirus expression vectors: A laboratory Manual, Oxford: Oxford University Press (1994).

Expressed poly-his tagged PRO polypeptide can then be purified, for example, by Ni^{2+}-chelate affinity chromatography as follows. Extracts are prepared from recombinant virus-infected Sf9 cells as described by Rupert *et al., Nature*, 362:175-179 (1993). Briefly, Sf9 cells are washed, resuspended in sonication buffer (25 mL Hepes, pH 7.9; 12.5 mM $MgCl_2$; 0.1 mM EDTA; 10% Glycerol; 0.1% NP-40; 0.4 M KCl), and sonicated twice for 20 seconds on ice. The sonicates are cleared by centrifugation, and the supernatant is diluted 50-fold in loading buffer (50 mM phosphate, 300 mM NaCl, 10% Glycerol, pH 7.8) and filtered through a 0.45 μm filter. A Ni^{2+}-NTA agarose column (commercially available from Qiagen) is prepared with a bed volume of 5 mL, washed with 25 mL of water and equilibrated with 25 mL of loading buffer. The filtered cell extract is loaded onto the column at 0.5 mL per minute. The column is washed to baseline A_{280} with loading buffer, at which point fraction collection is started. Next, the column is washed with a secondary wash buffer (50 mM phosphate; 300 mM NaCl, 10% Glycerol, pH 6.0), which elutes nonspecifically bound protein. After reaching A_{280} baseline again, the column is developed with a 0 to 500 mM Imidazole gradient in the secondary wash buffer. One mL fractions are collected and analyzed by SDS-PAGE and silver staining or western blot with Ni^{2+}-NTA-conjugated to alkaline phosphatase (Qiagen). Fractions containing the eluted His_{10}-tagged PRO polypeptide are pooled and dialyzed against loading buffer.

Alternatively, purification of the IgG tagged (or Fc tagged) PRO polypeptide can be performed using known chromatography techniques, including for instance, Protein A or protein G column chromatography.

PRO195, PRO526, PRO540, PRO846, PRO362, PRO363, PRO700, PRO707, PRO322, PRO719, PRO1083, PRO868, PRO866, PRO768, PRO788, PRO938, PRO827 and PRO1031 were successfully expressed in baculovirus infected Sf9 insect cells. While the expression was actually performed in a 0.5-2 L scale, it can be readily scaled up for larger (e.g. 8 L) preparations. The proteins were expressed as an IgG construct (immunoadhesin), in

which the protein extracellular region was fused to an IgG1 constant region sequence containing the hinge, CH2 and CH3 domains and/or in poly-His tagged forms.

For expression in baculovirus infected Sf9 cells, following PCR amplification, the respective coding sequences were subcloned into a baculovirus expression vector (pb.PH.IgG for IgG fusions and pb.PH.His.c for poly-His tagged proteins), and the vector and Baculogold® baculovirus DNA (Pharmingen) were co-transfected into 105 *Spodoptera frugiperda* ("Sf9") cells (ATCC CRL 1711), using Lipofectin (Gibco BRL). pb.PH.IgG and pb.PH.His are modifications of the commercially available baculovirus expression vector pVL1393 (Pharmingen), with modified polylinker regions to include the His or Fc tag sequences. The cells were grown in Hink's TNM-FH medium supplemented with 10% FBS (Hyclone). Cells were incubated for 5 days at 28°C. The supernatant was harvested and subsequently used for the first viral amplification by infecting Sf9 cells in Hink's TNM-FH medium supplemented with 10% FBS at an approximate multiplicity of infection (MOI) of 10. Cells were incubated for 3 days at 28°C. The supernatant was harvested and the expression of the constructs in the baculovirus expression vector was determined by batch binding of 1 ml of supernatant to 25 mL of Ni-NTA beads (QIAGEN) for histidine tagged proteins or Protein-A Sepharose CL-4B beads (Pharmacia) for IgG tagged proteins followed by SDS-PAGE analysis comparing to a known concentration of protein standard by Coomassie blue staining.

The first viral amplification supernatant was used to infect a spinner culture (500 ml) of Sf9 cells grown in ESF-921 medium (Expression Systems LLC) at an approximate MOI of 0.1. Cells were incubated for 3 days at 28°C. The supernatant was harvested and filtered. Batch binding and SDS-PAGE analysis was repeated, as necessary, until expression of the spinner culture was confirmed.

The conditioned medium from the transfected cells (0.5 to 3 L) was harvested by centrifugation to remove the cells and filtered through 0.22 micron filters. For the poly-His tagged constructs, the protein construct were purified using a Ni-NTA column (Qiagen). Before purification, imidazole was added to the conditioned media to a concentration of 5 mM. The conditioned media were pumped onto a 6 ml Ni-NTA column equilibrated in 20 mM Hepes, pH 7.4, buffer containing 0.3 M NaCl and 5 mM imidazole at a flow rate of 4-5 ml/min. at 4°C. After loading, the column was washed with additional equilibration buffer and the protein eluted with equilibration buffer containing 0.25 M imidazole. The highly purified protein was subsequently desalted into a storage buffer containing 10 mM Hepes, 0.14 M NaCl and 4% mannitol, pH 6.8, with a 25 ml G25 Superfine (Pharmacia) column and stored at -80°C.

Immunoadhesin (Fc containing) constructs of proteins were purified from the conditioned media as follows. The conditioned media were pumped onto a 5 ml Protein A column (Pharmacia) which had been equilibrated in 20 mM Na phosphate buffer, pH 6.8. After loading, the column was washed extensively with equilibration buffer before elution with 100 mM citric acid, pH 3.5. The eluted protein was immediately neutralized by collecting 1 ml fractions into tubes containing 275 mL of 1 M Tris buffer, pH 9. The highly purified protein was subsequently desalted into storage buffer as described above for the poly-His tagged proteins. The homogeneity of the proteins was verified by SDS polyacrylamide gel (PEG) electrophoresis and N-terminal amino acid sequencing by Edman degradation.

PRO181, PRO195, PRO200, PRO320, PRO237, PRO273, PRO285, PRO337, PRO526, PRO540, PRO846, PRO362, PRO363, PRO617, PRO322, PRO1083, PRO868, 768, PRO792, PRO788, PRO162, PRO1114, PRO827, PRO1075 and PRO1031 were successfully expressed in baculovirus infected Hi5 insect cells. While the expression was actually performed in a 0.5-2 L scale, it can be readily scaled up for larger (e.g. 8 L) preparations.

For expression in baculovirus-infected Hi5 insect cells, the PRO polypeptide-encoding DNA may be

amplified with suitable systems, such as Pfu (Stratagene), or fused upstream (5'-of) of an epitope tag contained with a baculovirus expression vector. Such epitope tags include poly-his tags and immunoglobulin tags (like Fc regions of IgG). A variety of plasmids may be employed, including plasmids derived from commercially available plasmids such as pVL1393 (Novagen). Briefly, the PRO polypeptide or the desired portion of the PRO polypeptide (such as the sequence encoding the extracellular domain of a transmembrane protein) is amplified by PCR with primers complementary to the 5' and 3' regions. The 5' primer may incorporate flanking (selected) restriction enzyme sites. The product is then digested with those selected restriction enzymes and subcloned into the expression vector. For example, derivatives of pVL1393 can include the Fc region of human IgG (pb.PH.IgG) or an 8 histidine (pb.PH.His) tag downstream (3'-of) the NAME sequence. Preferably, the vector construct is sequenced for confirmation.

Hi5 cells are grown to a confluency of 50% under the conditions of, 27°C, no CO2, NO pen/strep. For each 150 mm plate, 30 ug of pIE based vector containing PRO polypeptide is mixed with 1 ml Ex-Cell medium (Media: Ex-Cell 401 + 1/100 L-Glu JRH Biosciences #14401-78P (note: this media is light sensitive)), and in a separate tube, 100 ul of CellFectin (CellFECTIN (GibcoBRL #10362-010) (vortexed to mix)) is mixed with 1 ml of Ex-Cell medium. The two solutions are combined and allowed to incubate at room temperature for 15 minutes. 8 ml of Ex-Cell media is added to the 2ml of DNA/CellFECTIN mix and this is layered on Hi5 cells that have been washed once with Ex-Cell media. The plate is then incubated in darkness for 1 hour at room temperature. The DNA/CellFECTIN mix is then aspirated, and the cells are washed once with Ex-Cell to remove excess CellFECTIN . 30 ml of fresh Ex-Cell media is added and the cells are incubated for 3 days at 28°C. The supernatant is harvested and the expression of the PRO polypeptide in the baculovirus expression vector can be determined by batch binding of 1 ml of supernatent to 25 mL of Ni-NTA beads (QIAGEN) for histidine tagged proteins or Protein-A Sepharose CL-4B beads (Pharmacia) for IgG tagged proteins followed by SDS-PAGE analysis comparing to a known concentration of protein standard by Coomassie blue staining.

The conditioned media from the transfected cells (0.5 to 3 L) is harvested by centrifugation to remove the cells and filtered through 0.22 micron filters. For the poly-His tagged constructs, the protein comprising the PRO polypeptide is purified using a Ni-NTA column (Qiagen). Before purification, imidazole is added to the conditioned media to a concentration of 5 mM. The conditioned media is pumped onto a 6 ml Ni-NTA column equilibrated in 20 mM Hepes, pH 7.4, buffer containing 0.3 M NaCl and 5 mM imidazole at a flow rate of 4-5 ml/min. at 4°C. After loading, the column is washed with additional equilibration buffer and the protein eluted with equilibration buffer containing 0.25 M imidazole. The highly purified protein is subsequently deslated into a storage buffer containing 10 mM Hepes, 0.14 M NaCl and 4% mannitol, pH 6.8, with a 25 ml G25 Superfine (Pharmacia) column and stored at -80°C.

Immunoadhesin (Fc containing) constructs of proteins are purified from the conditioned media as follows. The conditioned media is pumped onto a 5 ml Protein A column (Pharmacia) which had been equilibrated in 20 mM Na phosphate buffer, pH 6.8. After loading, the column is washed extensively with equilibration buffer before elution with 100 mM citric acid, pH 3.5. The eluted protein is immediately neutralized by collecting 1 ml fractions into tubes containing 275 mL of 1 M Tris buffer, pH 9. The highly purified protein is subsequently desalted into storage buffer as described above for the poly-His tagged proteins. The homogeneity of PRO polypeptide can be assessed by SDS polyacrylamide gels and by N-terminal amino acid sequencing by Edman degradation and other analytical procedures as desired or necessary.

EXAMPLE 104: Preparation of Antibodies that Bind to PRO Polypeptides

This example illustrates preparation of monoclonal antibodies which can specifically bind to a PRO polypeptide.

Techniques for producing the monoclonal antibodies are known in the art and are described, for instance, in Goding, supra. Immunogens that may be employed include purified PRO polypeptide, fusion proteins containing the PRO polypeptide, and cells expressing recombinant PRO polypeptide on the cell surface. Selection of the immunogen can be made by the skilled artisan without undue experimentation.

Mice, such as Balb/c, are immunized with the PRO polypeptide immunogen emulsified in complete Freund's adjuvant and injected subcutaneously or intraperitoneally in an amount from 1-100 micrograms. Alternatively, the immunogen is emulsified in MPL-TDM adjuvant (Ribi Immunochemical Research, Hamilton, MT) and injected into the animal's hind foot pads. The immunized mice are then boosted 10 to 12 days later with additional immunogen emulsified in the selected adjuvant. Thereafter, for several weeks, the mice may also be boosted with additional immunization injections. Serum samples may be periodically obtained from the mice by retro-orbital bleeding for testing in ELISA assays to detect anti-PRO polypeptide antibodies.

After a suitable antibody titer has been detected, the animals "positive" for antibodies can be injected with a final intravenous injection of PRO polypeptide. Three to four days later, the mice are sacrificed and the spleen cells are harvested. The spleen cells are then fused (using 35% polyethylene glycol) to a selected murine myeloma cell line such as P3X63AgU.1, available from ATCC, No. CRL 1597. The fusions generate hybridoma cells which can then be plated in 96 well tissue culture plates containing HAT (hypoxanthine, aminopterin, and thymidine) medium to inhibit proliferation of non-fused cells, myeloma hybrids, and spleen cell hybrids.

The hybridoma cells will be screened in an ELISA for reactivity against the PRO polypeptide. Determination of "positive" hybridoma cells secreting the desired monoclonal antibodies against the PRO polypeptide is within the skill in the art.

The positive hybridoma cells can be injected intraperitoneally into syngeneic Balb/c mice to produce ascites containing the anti-PRO polypeptide monoclonal antibodies. Alternatively, the hybridoma cells can be grown in tissue culture flasks or roller bottles. Purification of the monoclonal antibodies produced in the ascites can be accomplished using ammonium sulfate precipitation, followed by gel exclusion chromatography. Alternatively, affinity chromatography based upon binding of antibody to protein A or protein G can be employed.

EXAMPLE 105: Chimeric PRO Polypeptides

PRO polypeptides may be expressed as chimeric proteins with one or more additional polypeptide domains added to facilitate protein purification. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS™ extension/affinity purification system (Immunex Corp., Seattle Wash.). The inclusion of a cleavable linker sequence such as Factor XA or enterokinase (Invitrogen, San Diego Calif.) between the purification domain and the PRO polypeptide sequence may be useful to facilitate expression of DNA encoding the PRO polypeptide.

EXAMPLE 106: Purification of PRO Polypeptides Using Specific Antibodies

Native or recombinant PRO polypeptides may be purified by a variety of standard techniques in the art of protein purification. For example, pro-PRO polypeptide, mature PRO polypeptide, or pre-PRO polypeptide is purified by immunoaffinity chromatography using antibodies specific for the PRO polypeptide of interest. In general, an immunoaffinity column is constructed by covalently coupling the anti-PRO polypeptide antibody to an activated chromatographic resin.

Polyclonal immunoglobulins are prepared from immune sera either by precipitation with ammonium sulfate or by purification on immobilized Protein A (Pharmacia LKB Biotechnology, Piscataway, N.J.). Likewise, monoclonal antibodies are prepared from mouse ascites fluid by ammonium sulfate precipitation or chromatography on immobilized Protein A. Partially purified immunoglobulin is covalently attached to a chromatographic resin such as CnBr-activated SEPHAROSE™ (Pharmacia LKB Biotechnology). The antibody is coupled to the resin, the resin is blocked, and the derivative resin is washed according to the manufacturer's instructions.

Such an immunoaffinity column is utilized in the purification of PRO polypeptide by preparing a fraction from cells containing PRO polypeptide in a soluble form. This preparation is derived by solubilization of the whole cell or of a subcellular fraction obtained via differential centrifugation by the addition of detergent or by other methods well known in the art. Alternatively, soluble PRO polypeptide containing a signal sequence may be secreted in useful quantity into the medium in which the cells are grown.

A soluble PRO polypeptide-containing preparation is passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of PRO polypeptide (*e.g.*, high ionic strength buffers in the presence of detergent). Then, the column is eluted under conditions that disrupt antibody/PRO polypeptide binding (*e.g.*, a low pH buffer such as approximately pH 2-3, or a high concentration of a chaotrope such as urea or thiocyanate ion), and PRO polypeptide is collected.

EXAMPLE 107: Drug Screening

This invention is particularly useful for screening compounds by using PRO polypeptides or binding fragment thereof in any of a variety of drug screening techniques. The PRO polypeptide or fragment employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the PRO polypeptide or fragment. Drugs are screened against such transformed cells in competitive binding assays. Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, the formation of complexes between PRO polypeptide or a fragment and the agent being tested. Alternatively, one can examine the diminution in complex formation between the PRO polypeptide and its target cell or target receptors caused by the agent being tested.

Thus, the present invention provides methods of screening for drugs or any other agents which can affect a PRO polypeptide-associated disease or disorder. These methods comprise contacting such an agent with an PRO polypeptide or fragment thereof and assaying (I) for the presence of a complex between the agent and the PRO polypeptide or fragment, or (ii) for the presence of a complex between the PRO polypeptide or fragment and the cell, by methods well known in the art. In such competitive binding assays, the PRO polypeptide or fragment is typically labeled. After suitable incubation, free PRO polypeptide or fragment is separated from that present in bound form, and the amount of free or uncomplexed label is a measure of the ability of the particular agent to bind to PRO

polypeptide or to interfere with the PRO polypeptide/cell complex.

Another technique for drug screening provides high throughput screening for compounds having suitable binding affinity to a polypeptide and is described in detail in WO 84/03564, published on September 13, 1984. Briefly stated, large numbers of different small peptide test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. As applied to a PRO polypeptide, the peptide test compounds are reacted with PRO polypeptide and washed. Bound PRO polypeptide is detected by methods well known in the art. Purified PRO polypeptide can also be coated directly onto plates for use in the aforementioned drug screening techniques. In addition, non-neutralizing antibodies can be used to capture the peptide and immobilize it on the solid support.

This invention also contemplates the use of competitive drug screening assays in which neutralizing antibodies capable of binding PRO polypeptide specifically compete with a test compound for binding to PRO polypeptide or fragments thereof. In this manner, the antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with PRO polypeptide.

EXAMPLE 108: Rational Drug Design

The goal of rational drug design is to produce structural analogs of biologically active polypeptide of interest (*i.e.*, a PRO polypeptide) or of small molecules with which they interact, *e.g.*, agonists, antagonists, or inhibitors. Any of these examples can be used to fashion drugs which are more active or stable forms of the PRO polypeptide or which enhance or interfere with the function of the PRO polypeptide *in vivo* (*c.f.*, Hodgson, Bio/Technology, 9: 19-21 (1991)).

In one approach, the three-dimensional structure of the PRO polypeptide, or of an PRO polypeptide-inhibitor complex, is determined by x-ray crystallography, by computer modeling or, most typically, by a combination of the two approaches. Both the shape and charges of the PRO polypeptide must be ascertained to elucidate the structure and to determine active site(s) of the molecule. Less often, useful information regarding the structure of the PRO polypeptide may be gained by modeling based on the structure of homologous proteins. In both cases, relevant structural information is used to design analogous PRO polypeptide-like molecules or to identify efficient inhibitors. Useful examples of rational drug design may include molecules which have improved activity or stability as shown by Braxton and Wells, Biochemistry, 31:7796-7801 (1992) or which act as inhibitors, agonists, or antagonists of native peptides as shown by Athauda *et al.*, J. Biochem., 113:742-746 (1993).

It is also possible to isolate a target-specific antibody, selected by functional assay, as described above, and then to solve its crystal structure. This approach, in principle, yields a pharmacore upon which subsequent drug design can be based. It is possible to bypass protein crystallography altogether by generating anti-idiotypic antibodies (anti-ids) to a functional, pharmacologically active antibody. As a mirror image of a mirror image, the binding site of the anti-ids would be expected to be an analog of the original receptor. The anti-id could then be used to identify and isolate peptides from banks of chemically or biologically produced peptides. The isolated peptides would then act as the pharmacore.

By virtue of the present invention, sufficient amounts of the PRO polypeptide may be made available to perform such analytical studies as X-ray crystallography. In addition, knowledge of the PRO polypeptide amino acid sequence provided herein will provide guidance to those employing computer modeling techniques in place of or in addition to x-ray crystallography.

EXAMPLE 109: Ability of PRO Polypeptides to Inhibit Vascular Endothelial Growth Factor (VEGF) Stimulated Proliferation of Endothelial Cell Growth

The ability of various PRO polypeptides to inhibit VEGF stimulated proliferation of endothelial cells was tested. Specifically, bovine adrenal cortical capillary endothelial (ACE) cells (from primary culture, maximum 12-14 passages) were plated on 96-well microtiter plates (Amersham Life Science) at a density of 500 cells/well per 100 μL in low glucose DMEM, 10% calf serum, 2 mM glutamine, 1x pen/strept and fungizone, supplemented with 3 ng/mL VEGF. Controls were plated the same way but some did not include VEGF. A test sample of the PRO polypeptide of interest was added in a 100 μl volume for a 200 μl final volume. Cells were incubated for 6-7 days at 37°C. The media was aspirated and the cells washed 1x with PBS. An acid phosphatase reaction mixture (100 μL, 0.1M sodium acetate, pH 5.5, 0.1% Triton-100, 10 mM p-nitrophenyl phosphate) was added. After incubation for 2 hours at 37°C, the reaction was stopped by addition of 10 μl 1N NaOH. OD was measured on microtiter plate reader at 405 nm. Controls were no cells, cells alone, cells + FGF (5 ng/mL), cells + VEGF (3 ng/mL), cells + VEGF (3 ng/ml) + TGF-β (1 ng/ml), and cells + VEGF (3ng/mL) + LIF (5 ng/mL). (TGF-β at a 1 ng/ml concentration is known to block 70-90% of VEGF stimulated cell proliferation.)

The results were assessed by calculating the percentage inhibition of VEGF (3 ng/ml) stimulated cells proliferation, determined by measuring acid phosphatase activity at OD405 nm, (1) relative to cells without stimulation, and (2) relative to the reference TGF-β inhibition of VEGF stimulated activity. The results are indicative of the utility of the PRO polypeptides in cancer therapy and specifically in inhibiting tumor angiogenesis.

The PRO polypeptides demonstrated as being capable of inhibiting VEGF stimulated proliferation of endothelial cell growth at various concentrations include PRO200 and PRO320.

EXAMPLE 110: Retinal Neuron Survival

This example demonstrates that various PRO polypeptides have efficacy in enhancing the survival of retinal neuron cells.

Sprague Dawley rat pups at postnatal day 7 (mixed population: glia and retinal neuronal types) are killed by decapitation following CO_2 anesthesia and the eyes are removed under sterile conditions. The neural retina is dissected away from the pigment epithelium and other ocular tissue and then dissociated into a single cell suspension using 0.25% trypsin in Ca^{2+}, Mg^{2+}-free PBS. The retinas are incubated at 37°C for 7-10 minutes after which the trypsin is inactivated by adding 1 ml soybean trypsin inhibitor. The cells are plated at 100,000 cells per well in 96 well plates in DMEM/F12 supplemented with N2 and with or without the specific test PRO polypeptide. Cells for all experiments are grown at 37°C in a water saturated atmosphere of 5% CO_2. After 2-3 days in culture, cells are stained with calcein AM then fixed using 4% paraformaldehyde and stained with DAPI for determination of total cell count. The total cells (fluorescent) are quantified at 20X objective magnification using CCD camera and NIH image software for MacIntosh. Fields in the well are chosen at random.

The effect of various concentration of PRO polypeptides is calculated by dividing the total number of calcein AM positive cells at 2-3 days in culture by the total number of DAPI-labeled cells at 2-3 days in culture. Anything above 30% survival is considered positive. The following PRO polypeptides were positive in this assay: PRO200, PRO540, PRO846 and PRO617.

EXAMPLE 111: Rod Photoreceptor Survival

This example demonstrates that various PRO polypeptides have efficacy in enhancing the survival of rod photoreceptor cells.

Sprague Dawley rat pups at 7 day postnatal (mixed population: glia and retinal neuronal cell types) are killed by decapitation following CO_2 anesthesis and the eyes are removed under sterile conditions. The neural retina is dissected away form the pigment epithelium and other ocular tissue and then dissociated into a single cell suspension using 0.25% trypsin in Ca^{2+}, Mg^{2+}-free PBS. The retinas are incubated at 37°C for 7-10 minutes after which the trypsin is inactivated by adding 1 ml soybean trypsin inhibitor. The cells are plated at 100,000 cells per well in 96 well plates in DMEM/F12 supplemented with N2 and with or without the specific test PRO polypeptide. Cells for all experiments are grown at 37°C in a water saturated atmosphere of 5% CO2. After 2-3 days in culture, cells are fixed using 4% paraformaldehyde, and then stained using CellTracker Green CMFDA. Rho 4D2 (ascites or IgG 1:100), a monoclonal antibody directed towards the visual pigment rhodopsin is used to detect rod photoreceptor cells by indirect immunofluorescence. The results are reported as % survival: total number of calcein/CellTracker - rhodopsin positive cells at 2-3 days in culture, divided by the total number of rhodopsin positive cells at time 2-3 days in culture. The total cells (fluorescent) are quantified at 20x objective magnification using a CCD camera and NIH image software for MacIntosh. Fields in the well are chosen at random.

With regard to the effect of various concentration of PRO polypeptides, anything above 10% survival is considered positive. The following PRO polypeptides tested positive in this assay: PRO200, PRO540, PRO846 and PRO617.

EXAMPLE 112: Ability of PRO Polypeptides to Stimulate the Release of Proteoglycans from Cartilage

The ability of various PRO polypeptides to stimulate the release of proteoglycans from cartilage tissue was tested as follows.

The metacarphophalangeal joint of 4-6 month old pigs was aseptically dissected, and articular cartilage was removed by free hand slicing being careful to avoid the underlying bone. The cartilage was minced and cultured in bulk for 24 hours in a humidified atmosphere of 95% air, 5% CO_2 in serum free (SF) media (DME/F12 1:1) woth 0.1% BSA and 100U/ml penicillin and 100μg/ml streptomycin. After washing three times, approximately 100 mg of articular cartilage was aliquoted into micronics tubes and incubated for an additional 24 hours in the above SF media. PRO polypeptides were then added at 1% either alone or in combination with 18 ng/ml interleukin-1α, a known stimulator of proteoglycan release from cartilage tissue. The supernatant was then harvested and assayed for the amount of proteoglycans using the 1,9-dimethyl-methylene blue (DMB) colorimetric assay (Farndale and Buttle, Biochem. Biophys. Acta 883:173-177 (1985)). A positive result in this assay indicates that the test polypeptide will find use, for example, in the treatment of sports-related joint problems, articular cartilage defects, osteoarthritis or rheumatoid arthritis.

When PRO200 polypeptides were tested in the above assay, the polypeptides demonstrated a marked ability to stimulate release of proteoglycans from cartilage tissue both basally and after stimulation with interleukin-1α and at 24 and 72 hours after treatment, thereby indicating that PRO200 polypeptides are useful for stimulating proteoglycan release from cartilage tissue.

EXAMPLE 113: In Vitro Antiproliferative Assay

The antiproliferative activity of various PRO polypeptides was determined in the investigational, disease-oriented *in vitro* anti-cancer drug discovery assay of the National Cancer Institute (NCI), using a sulforhodamine B (SRB) dye binding assay essentially as described by Skehan et al., *J. Natl. Cancer Inst.* 82:1107-1112 (1990). The 60 tumor cell lines employed in this study ("the NCI panel"), as well as conditions for their maintenance and culture *in vitro* have been described by Monks et al., *J. Natl. Cancer Inst.* 83:757-766 (1991). The purpose of this screen is to initially evaluate the cytotoxic and/or cytostatic activity of the test compounds against different types of tumors (Monks et al., *supra*; Boyd, *Cancer: Princ. Pract. Oncol. Update* 3(10):1-12 [1989]).

Cells from approximately 60 human tumor cell lines were harvested with trypsin/EDTA (Gibco), washed once, resuspended in IMEM and their viability was determined. The cell suspensions were added by pipet (100 μL volume) into separate 96-well microtiter plates. The cell density for the 6-day incubation was less than for the 2-day incubation to prevent overgrowth. Inoculates were allowed a preincubation period of 24 hours at 37°C for stabilization. Dilutions at twice the intended test concentration were added at time zero in 100 μL aliquots to the microtiter plate wells (1:2 dilution). Test compounds were evaluated at five half-log dilutions (1000 to 100,000-fold). Incubations took place for two days and six days in a 5% CO_2 atmosphere and 100% humidity.

After incubation, the medium was removed and the cells were fixed in 0.1 ml of 10% trichloroacetic acid at 40°C. The plates were rinsed five times with deionized water, dried, stained for 30 minutes with 0.1 ml of 0.4% sulforhodamine B dye (Sigma) dissolved in 1% acetic acid, rinsed four times with 1% acetic acid to remove unbound dye, dried, and the stain was extracted for five minutes with 0.1 ml of 10 mM Tris base [tris(hydroxymethyl)aminomethane], pH 10.5. The absorbance (OD) of sulforhodamine B at 492 nm was measured using a computer-interfaced, 96-well microtiter plate reader.

A test sample is considered positive if it shows at least 50% growth inhibitory effect at one or more concentrations. The following PRO polypeptides gave positive results in at least one tumor cell line: PRO181, PRO237, PRO526, PRO362 and PRO866.

EXAMPLE 114: Gene Amplification

This example shows that genes encoding various PRO polypeptides are amplified in the genome of certain human cancers. Amplification is associated with overexpression of the gene product, indicating that the PRO polypeptide is a useful target for therapeutic intervention in certain cancers such as colon, lung and other cancers. Therapeutic agent may take the form of antagonists of PRO polypeptide-encoding genes, for example, murine-human chimeric, humanized or human antibodies against the PRO polypeptide.

The starting material for the screen was genomic DNA isolated from a variety cancers. The DNA is quantitated precisely, e.g., fluorometrically. As a negative control, DNA was isolated from the cells of ten normal healthy individuals which was pooled and used as assay controls for the gene copy in healthy individuals (NorHu).

The 5' nuclease assay (for example, TaqMan™) and real-time quantitative PCR (for example, ABI Prizm 7700 Sequence Detection System™ (Perkin Elmer, Applied Biosystems Division, Foster City, CA)), were used to find genes potentially amplified in certain cancers. The results were used to determine whether the DNA encoding the PRO polypeptide is over-represented in any of the lung and colon cancers that were screened. The result was reported in Delta CT units. One unit corresponds 1 PCR cycle or approximately a 2-fold amplification relative to normal, two units corresponds to 4-fold, 3 units to 8-fold and so on. Quantitation was obtained using primers and

a Taqman™ fluorescent derived from the PRO polypeptide-encoding gene. Regions of the PRO polypeptide which are most likely to contain unique nucleic acid sequences and which are least likely to have spliced out introns are preferred for the primer derivation, e.g., 3'-untranslated region.

The 5' nuclease assay reaction is a fluorescent PCR-based technique which makes use of the 5' exonuclease activity of Taq DNA polymerase enzyme to monitor amplification in real time. Two oligonucleotide primers are used to generate an amplicon typical of a PCR reaction. A third oligonucleotide, or probe, is designed to detect nucleotide sequence located between the two PCR primers. The probe is non-extendible by Taq DNA polymerase enzyme, and is labeled with a reporter fluorescent dye and a quencher fluorescent dye. Any laser-induced emission from the reporter dye is quenched by the quenching dye when the two dyes are located close together as they are on the probe. During the amplification reaction, the probe is cleaved by the Taq DNA polymerase enzyme in a template-dependent manner. The resultant probe fragments disassociate in solution, and signal from the released reporter dye is free from the quenching effect of the second fluorophore. One molecule of reporter dye is liberated for each new molecule synthesized, and detection of the unquenched reporter dye provides the basis for quantitative interpretation of the data.

The 5' nuclease procedure is run on a real-time quantitative PCR device such as the ABI Prism 7700TM Sequence Detection. The system consists of a thermocyler, laser, charge-coupled device (CCD) camera and computer. The system amplifies samples in a 96-well format on a thermocycler. During amplification, laser-induced fluorescent signal is collected in real-time through fiber optics cables for all 96 wells, and detected at the CCD. The system includes software for running the intrument and for analyzing the data.

5' Nuclease assay data are initially expressed as Ct, or the threshold cycle. This is defined as the cycle at which the reporter signal accumulates above the background level of fluorescence. The Ct values are used as quantitative measurement of the relative number of starting copies of a particular target sequence in a nucleic acid sample.

Genes encoding the following PRO polypeptides were found to be amplified in the above assay: PRO213-1, PRO237, PRO324, PRO351, PRO362, PRO853, PRO615, PRO531, PRO618, PRO772, PRO703, PRO474, PRO1017 and PRO792.

EXAMPLE 115: Induction of c-fos in Endothelial Cells

Human venous umbilical vein endothelial cells (HUVEC, Cell Systems) in growth media (50:50 without glycine, 1% glutamine, 10mM Hepes, 10% FBS, 10 ng/ml bFGF), were plated on 96-well microtiter plates at a cell density of $1x10^4$ cells/well. The day after plating, the cells were starved by removing the growth media and treating the cells with 100 µl/well test samples and controls (positive control: growth media; negative control: Protein 32). The cells were incubated for 30 minutes at 37°C, in 5% CO_2. The samples were removed, and the first part of the bDNA kit protocol (Chiron Diagnostics, cat. #6005-037) was followed.

Briefly, the amounts of the TM Lysis Buffer and Probes needed for the tests were calculated based on information provided by the manufacturer. The appropriate amounts of thawed Probes were added to the TM Lysis Buffer. The Capture Hybridization Buffer was warmed to room temperature. The bDNA strips were set up in the metal strip holders, and 100 µl of Capture Hybridization Buffer were added to each b-DNA well needed, followed by incubation for at least 30 minutes. The test plates with the cells were removed from the incubator, and the media was gently removed using the vacuum manifold. 100µl of Lysis Hybridization Buffer with Probes were quickly pipetted into each well of the microtiter plates. The plates were then incubated at 55°C for 15 minutes. Upon

removal from the incubator, the plates were placed on the vortex mixer with the microtiter adapter head and vortex on the #2 setting for one minute. 80 μl of the lysate were removed and added to the bDNA wells containing the Capture Hybridization Buffer, and pipetted up and down to mix. The plates were incubated at 53°C for at least 16 hours.

On the next day, the second part of the bDNA kit protocol was followed. Specifically, the Plates were removed from the incubator and placed on the bench to cool for 10 minutes. The volumes of additions needed were calculated based upon information provided by the manufacturer. An Amplifier Working Solution was prepared by making a 1:100 dilution of the Amplifier Concentrate (20 fm/μl) in AL Hybridization Buffer. The hybridization mixture was removed from the plates and washed twice with Wash A. 50 μl of Amplifier Working Solution were added to each well and the wells were incubated at 53°C for 30 minutes. The plates were then removed from the incubator and allowed to cool for 10 minutes. The Label Probe Working Solution was prepared by making a 1:100 dilution of Label Concentrate (40 pmoles/μl) in AL Hybridization Buffer. After the 10 minutes cool down period, the amplifier hybridization mixture was removed and the plates washed twice with Wash A. 50μl of Label Probe Working Solution were added to each well and the wells were incubated at 53°C for 15 minutes. After cooling for 10 minutes, the Substrate was warmed to room temperature. Upon addition of 3 μl of Substrate Enhancer to each ml of Substrate needed for the assay, the plates were allowed to cool for 10 minutes, the label hybridization mixture was removed, and the plates were washed twice with Wash A and three-times with Wash D. 50μl of the Substrate Solution with Enhancer were added to each well. The plates were incubated for 30 minutes at 37°C and RLU read in an appropriate luminometer.

The replicates were averaged and the coefficient of variation was determined. The measure of activity of the fold increase over the Protein 32 (buffer control) value indicated by chemoluminescence units (RLU). Samples which showed an at least two-fold value over the Protein 32 value were considered positive. PRO938 was positive in the above assay.

EXAMPLE 116: Proliferation of Rat Utricular Supporting Cells

In an effort to identify PRO polypeptides that act as potent mitogens for inner ear supporting cells which are hair cell progenitors (related to auditory hair cell regeneration), various PRO polypeptides were tested in the following assay.

The rat utricular epithelial cell line (UEC-4 cells) are aliquoted into 96 well plates with a density of 3000 cells/well in 200 μl of serum-containing medium at 33°C. After overnight, the cultures are switched to serum-free medium at 37°C and the PRO polypeptide samples are added at various dilutions. After 24h incubation, ^{3}H-thymidine (1 μCi/well) is aded to the cultures for an additional 24h. The cells are then harvested using a Tomtec cell harvester. Because the epithelial cells are grown on a polylysine substrate, trypsin (1 mg/ml) is added to the culture wells for 30 min at 37°C to lift the cells before cell harvest. Cpm/well are counted with a matrix 9600 gas counter (Packard Instrument Company, Downers Grove, IL). Data is collected from 3 culture wells from each of the experimental groups and expressed as mean ± SEM. A two-tailed, unpaired t-test is used for statistical analysis, as compared to the control group (treatment with TGF-α).

Average cpm counts which are at least 30% higher than the control values are considered positive for the assay. The following PRO polypeptides were positive in this assay: PRO337, PRO363 and PRO1012.

EXAMPLE 117: *In situ* Hybridization

In situ hybridization is a powerful and versatile technique for the detection and localization of nucleic acid sequences within cell or tissue preparations. It may be useful, for example, to identify sites of gene expression, analyze the tissue distribution of transcription, identify and localize viral infection, follow changes in specific mRNA synthesis and aid in chromosome mapping.

In situ hybridization was performed following an optimized version of the protocol by Lu and Gillett, Cell Vision 1:169-176 (1994), using PCR-generated ^{33}P-labeled riboprobes. Briefly, formalin-fixed, paraffin-embedded human tissues were sectioned, deparaffinized, deproteinated in proteinase K (20 g/ml) for 15 minutes at 37°C, and further processed for *in situ* hybridization as described by Lu and Gillett, *supra*. A [^{33}-P] UTP-labeled antisense riboprobe was generated from a PCR product and hybridized at 55°C overnight. The slides were dipped in Kodak NTB2 nuclear track emulsion and exposed for 4 weeks.

^{33}P-Riboprobe synthesis

6.0 μl (125 mCi) of ^{33}P-UTP (Amersham BF 1002, SA<2000 Ci/mmol) were speed vac dried. To each tube containing dried ^{33}P-UTP, the following ingredients were added:

2.0 μl 5x transcription buffer
1.0 μl DTT (100 mM)
2.0 μl NTP mix (2.5 mM : 10 μ; each of 10 mM GTP, CTP & ATP + 10 μl H_2O)
1.0 μl UTP (50 μM)
1.0 μl Rnasin
1.0 μl DNA template (1μg)
1.0 μl H_2O
1.0 μl RNA polymerase (for PCR products T3 = AS, T7 = S, usually)

The tubes were incubated at 37°C for one hour. 1.0 μl RQ1 DNase were added, followed by incubation at 37°C for 15 minutes. 90 μl TE (10 mM Tris pH 7.6/1mM EDTA pH 8.0) were added, and the mixture was pipetted onto DE81 paper. The remaining solution was loaded in a Microcon-50 ultrafiltration unit, and spun using program 10 (6 minutes). The filtration unit was inverted over a second tube and spun using program 2 (3 minutes). After the final recovery spin, 100 μl TE were added. 1 μl of the final product was pipetted on DE81 paper and counted in 6 ml of Biofluor II.

The probe was run on a TBE/urea gel. 1-3 μl of the probe or 5 μl of RNA Mrk III were added to 3 μl of loading buffer. After heating on a 95°C heat block for three minutes, the gel was immediately placed on ice. The wells of gel were flushed, the sample loaded, and run at 180-250 volts for 45 minutes. The gel was wrapped in saran wrap and exposed to XAR film with an intensifying screen in -70°C freezer one hour to overnight.

^{33}P-Hybridization

A. Pretreatment of frozen sections

The slides were removed from the freezer, placed on aluminium trays and thawed at room temperature for 5 minutes. The trays were placed in 55°C incubator for five minutes to reduce condensation. The slides were fixed for 10 minutes in 4% paraformaldehyde on ice in the fume hood, and washed in 0.5 x SSC for 5 minutes, at room temperature (25 ml 20 x SSC + 975 ml SQ H_2O). After deproteination in 0.5 μg/ml proteinase K for 10 minutes at 37°C (12.5 μl of 10 mg/ml stock in 250 ml prewarmed RNase-free RNAse buffer), the sections were washed in 0.5 x SSC for 10 minutes at room temperature. The sections were dehydrated in 70%, 95%, 100% ethanol, 2

minutes each.

B. Pretreatment of paraffin-embedded sections

The slides were deparaffinized, placed in SQ H_2O, and rinsed twice in 2 x SSC at room temperature, for 5 minutes each time. The sections were deproteinated in 20 μg/ml proteinase K (500 μl of 10 mg/ml in 250 ml RNase-free RNase buffer; 37°C, 15 minutes) - human embryo, or 8 x proteinase K (100 μl in 250 ml Rnase buffer, 37°C, 30 minutes) - formalin tissues. Subsequent rinsing in 0.5 x SSC and dehydration were performed as described above.

C. Prehybridization

The slides were laid out in a plastic box lined with Box buffer (4 x SSC, 50% formamide) - saturated filter paper. The tissue was covered with 50 μl of hybridization buffer (3.75g Dextran Sulfate + 6 ml SQ H_2O), vortexed and heated in the microwave for 2 minutes with the cap loosened. After cooling on ice, 18.75 ml formamide, 3.75 ml 20 x SSC and 9 ml SQ H_2O were added, the tissue was vortexed well, and incubated at 42°C for 1-4 hours.

D. Hybridization

1.0 x 10^6 cpm probe and 1.0 μl tRNA (50 mg/ml stock) per slide were heated at 95°C for 3 minutes. The slides were cooled on ice, and 48 μl hybridization buffer were added per slide. After vortexing, 50 μl ^{33}P mix were added to 50 μl prehybridization on slide. The slides were incubated overnight at 55°C.

E. Washes

Washing was done 2 x 10 minutes with 2xSSC, EDTA at room temperature (400 ml 20 x SSC + 16 ml 0.25M EDTA, V_f=4L), followed by RNaseA treatment at 37°C for 30 minutes (500 μl of 10 mg/ml in 250 ml Rnase buffer = 20 μg/ml). The slides were washed 2 x 10 minutes with 2 x SSC, EDTA at room temperature. The stringency wash conditions were as follows: 2 hours at 55°C, 0.1 x SSC, EDTA (20 ml 20 x SSC + 16 ml EDTA, V_f=4L).

F. Oligonucleotides

In situ analysis was performed on a variety of DNA sequences disclosed herein. The oligonucleotides employed for these analyses were derived from the nucleotide sequences disclosed herein and generally range from about 40 to 55 nucleotides in length.

G. Results

In situ analysis was performed on a variety of DNA sequences disclosed herein. The results from these analyses are as follows.

(1) DNA29101-1122 (PRO200)

Fetal: Lower limb expression in developing lower limb bones at the edge of the cartilagenous anlage (i.e. around the outside edge); in developing tendons, in vascular smooth muscle and in cells embracing developing skeletal muscle myocytes and myotubes. Expression also observed at the epiphyseal growth plate. Lymph node expression in marginal sinus of developing lymph nodes. Thymus expression in the subcapsular region of the thymic cortex, possibly representing either the subcapsular epithelial cells or the proliferating, double negative, thymocytes that are found in this region. Spleen is negative. Trachea expression in smooth muscle. Brain (cerebral cortex) focal expression in cortical neurones. Spinal cord negative. Small intestine expression in smooth muscle. Thyroid - generalized expression over thyroid epithelium. Adrenal is negative. Liver expression in ductal plate cells. Stomach expression in mural smooth muscle. Fetal skin expression in basal layer of squamous epithelium. Placenta expression in interstitial cells in trophoblastic villi. Cord expression in wall of arteries and vein.

Comments: Expression pattern suggests that PRO200 may be involved in cell differentiation/proliferation. High expression was observed at the following additional sites: Chimp ovary - granulosa cells of maturing follicles, lower intensity signal observed over thecal cells. Chimp parathyroid - high expression over chief cells. Human fetal testis - moderate expression over stromal cells surrounding developing tubules. Human fetal lung - high expression over chondrocytes in developing bronchial tree, and low level expression over branching bronchial epithelium. Specific expression was not observed over the renal cell, gastric and colonic carcinomas. Fetal tissues examined (E12-E16 weeks) include: placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis and lower limb. Adult tissues examined: liver, kidney, adrenal, myocardium, aorta, spleen, lymph node, pancreas, lung, skin, cerebral cortex (rm), hippocampus(rm), cerebellum(rm), penis, eye, bladder, stomach, gastric carcinoma, colon, colonic carcinoma and chondrosarcoma. Acetominophen induced liver injury and hepatic cirrhosis.

(2) DNA30867-1335 (PRO218)

Low level expression over numerous epithelia including fetal small intestine, fetal thyroid, chimp gastric epithelium. Expression also seen over malignant cells in a renal cell carcinoma. Expression in fetal brain, over cortex. The distribution does not suggest an obvious function. Human fetal tissues examined (E12-E16 weeks) include: placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis and lower limb. Adult human tissues examined: kidney (normal and end-stage), bladder, adrenal, spleen, lymph node, pancreas, lung, skin, eye (inc. retina), colon, bladder, liver (normal, cirrhotic, acute failure), heart, clear cell carcinoma of kidney, gastric adenocarcinoma, colorectal carcinoma. Non-human primate tissues examined: Chimp tissues: salivary gland, stomach, thyroid, parathyroid, tongue, thymus, ovary, lymph node, peripheral nerve. Rhesus Monkey tissues: cerebral cortex, hippocampus, cerebellum, penis.

(3) DNA40021-1154 (PRO285)

Low levels of expression observed in the placenta and over hematopoietic cells in the mouse fetal liver. No expression was detected in either human fetal, adult or chimp lymph node and no expression was detected in human fetal or human adult spleen. Fetal tissues examined (E12-E16 weeks) include: placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis and lower limb. Adult tissues examined: liver, kidney, adrenal, myocardium, aorta, spleen, lymph node, pancreas, lung, skin, cerebral cortex (rm), hippocampus(rm), cerebellum(rm), brain infarct (human), cerebritis (human),penis, eye, bladder, stomach, gastric carcinoma, colon, colonic carcinoma, thyroid (chimp), parathyroid (chimp) ovary (chimp) and chondrosarcoma. Acetominophen induced liver injury and hepatic cirrhosis.

(4) DNA39523-1192 (PRO273)

Expression over epithelium of mouse embryo skin as well as over basal epithelium and dermis of human fetal skin. Basal epithelial pegs of the squamous mucosa of the chimp tongue are also positive. Expression over a subset of cells in developing glomeruli of fetal kidney, adult renal tubules, and over "thyroidized" epithelium in end-stage renal disease, low expression in a renal cell carcinoma, probably over the epithelial cells. Low level expression

over stromal cells in fetal lung. Expression over stromal cells in the apical portion of gastric glands. High expression in the lamina propria of the fetal small intestinal villi, normal colonic mucosa and over stromal cells in a colonic carcinoma. Strong expression over benign connective tissue cells in the hylanized stroma of a sarcoma. Expression over stromal cells in the placental villi and the splenic red pulp. In the brain, expression over cortical neurones. Connective tissue surrounding developing bones and over nerve sheath cells in the fetus. Fetal tissues examined (E12-E16 weeks) include: placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis and lower limb. Adult tissues examined: liver, kidney, adrenal, myocardium, aorta, spleen, lymph node, pancreas, lung, skin, cerebral cortex (rm), hippocampus(rm), eye, stomach, gastric carcinoma, colon, colonic carcinoma, thyroid (chimp), parathyroid (chimp) ovary (chimp) and chondrosarcoma. Acetominophen induced liver injury and hepatic cirrhosis.

Expression was present in many cells in the outer layers (I and II) of the monkey cerebral cortex. A small subset of cells in the deeper cortical layers also expressed mRNA for this chemokine homolog. Scattered cells within the molecular layers of the hippocampus and bordering the inner edge of the dentate gyrus contained chemokine homolog mRNA. No expression was detected within the cerebellar cortex. Chemokine homolog expression is not observed in infarcted brain, where cell death has occurred in the regions where the chemokine homolog normally is expressed. This probe could possibly serve as a marker of a subset of neurons of outer layers of the cerebral cortex and could possibly reveal neuronal migration disorders. Abnormal neuronal migration is a possible cause of some seizure disorders and schizophrenia. In order to gain a better appreciation of the distribution of this mRNA we will test whether the probe will cross-hybridize with mouse brain tissue.

Also shows intriguing and specific patterns of hybridization within postnatal day (P)10 and adult mouse brains. In one sagittal section of P10 mouse brain, strong signal was observed scattered within the molecular layer of the hippocampus and inner edges of the dentate gyrus. Cells in the presubiculum were moderately labeled; the signal extended in a strong band through outer layers of the retrosplenial cortes to the occipital cortex, where the signal diminished to background levels. A small set of positive neurons were detected in deeper regions of P10 motor cortex; neurons in outer layers of P10 cortex did not exhibit signal above background levels. Moderate hybridization signal was also detected in the inferior colliculus. Chemokine homolog signal in the adult mouse brain was evaluated in three coronal sections at different levels. Strong signal was detected in the septum and in scattered neurons in the pontine nuclei and motor root of the trigeminal nerve; moderate signal was seen in the molecular layers of the hippocampus and outer layers of the retrosplenial cortex.

(5) DNA39979-1213 (PRO296)

Widespread expression in fetal in adult tissues. Expressed in a variety of fetal and adult epithelia, skeletal and cardiac muscle, developing (including retina) and adult CNS, thymic epithelium, placental villi, hepatocytes in cirrhotic and acetaminophen induced toxicity. Highly expressed in hypertrophic chondrocytes in developing skeletal system.The overall expression pattern, while not completely ovelapping (not expressed in glomeruli, more widely expressed in CNS), is not disimilar to VEGF. A possible role in angiogenesis should therefore be considered. Human fetal tissues examined (E12-E16 weeks) include: placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, great vessels, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis, testis and lower limb. Adult human tissues examined: kidney (normal and end-stage), adrenal, spleen, lymph node, pancreas, lung, eye (inc. retina), bladder, liver (normal, cirrhotic, acute failure). Non-human primate tissues

examined: Chimp tissues: adrenal. Rhesus Monkey tissues: cerebral cortex, hippocampus, cerebellum.

(6) DNA52594-1270 (PRO868)

Expression over neuronal cells in fetal dorsal root ganglia, spinal cord, developing enteric neurons, cortical neurons. Low level expression also seen in placental trophoblast. In adult tissues the only site where notable expression was observed was the normal adult prostate; as such it may represent a possible prostate cell surface receptor target antigen. Studies to further characterize the expression in adult tissues seem warranted. Low level expression also observed in a liposarcoma. Fetal tissues examined (E12-E16 weeks) include: placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis and lower limb. Adult human tissues examined: liver, kidney, adrenal, myocardium, aorta, spleen, lung, skin, chondrosarcoma, eye, stomach, gastric carcinoma, colon, colonic carcinoma, renal cell carcinoma, prostate, bladder mucosa and gall bladder. Acetominophen induced liver injury and hepatic cirrhosis. Rhesus tissues examined: cerebral cortex (rm), hippocampus(rm), cerebellum. Chimp tissues examined: thyroid, parathyroid, ovary, nerve, tongue, thymus, adrenal, gastric mucosa and salivary gland. WIG-1(WISP-1), WIG-2 (WISP-2) and WIG-5 (WISP-3) expression in human breast carcinoma and normal breast tissue, Wig-2 in lung carcinoma, and Wig-5 in colon carcinoma.

(7) DNA64907-1163 (PRO1330)

In human fetal tissues there was strong specific expression over artrerial, venous, capillary and sinusoidal endothelium in all tissues examined, except for fetal brain. In normal adult tissues expression was low to absent, but when present appeared expression was confined to the vasculature. Highest expression in adult tissues was observed regionally in vessels running within the white matter of rhesus brain - the significance of this pattern is unclear. Elevated expression observed in vasculature of many inflamed and diseased tissues, including tumor vasculature. In some of these tissues it was unclear if expression was soley confined to vascular endothelium. In the 15 lung tumors examined no expression was seen over the malignant epithelium, however, vascular expression was observed in many of the tumors and adjacent lung tissue. Moderate, apparently non-specific background, was seen with this probe over hyalinised collagen and sites of tissue necrosis. In the abscence of a sense control, however, it is not possible to be absolutely certain that all of this signal is non-specific. Some signal, also thought to be non-specific, was seen over the chimp gastric mucosa, transitional cell epithelium of human adult bladder and fetal retina.

(8) DNA49624-1279 (PRO545)

Expression of the ADAM family molecule, ADAM 12 (DNA49624-1279) observed in normal human lung, lung tumor, normal colon and colon carcinoma.

(9) DNA59294-1381 (PRO1031)

The expression of this IL17 homologue was evaluated in a panel consisting of normal adult and fetal tissues and tissues with inflammation, predominantly chronic lymphocytic inflammation. This panel is designed to specifically evaluate the expression pattern in immune mediated inflammatory disease of novel proteins that modulate T lymphocyte function (stimulatory or inhibitory). This protein when expressed as an Ig-fusion protein was immunostimulatory in a dose dependent fashion in the human mixed lymphocyte reaction (MLR); it caused a 285%

and 147% increase above the baseline stimulation index when utilized at two different concentrations (1.0% and 0.1% of a 560 nM stock). Summary: expression was restricted to muscle, certain types of smooth muscle in the adult and in skeletal and smooth muscle in the human fetus. Expression in adult human was in smooth muscle of tubular organs evaluated including colon and gall bladder. There no expression in the smooth muscle of vessels or bronchi. No adult human skeletal muscle was evaluated. In fetal tissues there was moderate to high diffuse expression in skeletal muscle the axial skeleton and limbs. There was weak expression in the smooth muscle of the intestinal wall but no expression in cardiac muscle. Adult human tissues with expression: Colon. there was low level diffuse expression in the smooth muscle (tunica muscularis) in 5 specimens with chronic inflammatory bowel disease. Gall bladder: there was weak to low level expression in the smooth muscle of the gall bladder. Fetal human tissues with expression: there was moderate diffuse expression in skeletal muscle and weak tolow expression in smooth muscle; there was no expression in fetal heart or any other fetal organ including liver, spleen, CNS, kidney, gut, lung. Human tissues with no expression: lung with chronic granulomatous inflammation and chronic bronchitis (5 patients), peripheral nerve, prostate, heart, placenta, liver (disease multiblock), brain (cerbrum and cerebellum), tonsil (reactive hyperplasia), peripheral lymph node, thymus.

(10) DNA45416-1251 (PRO362)

The expression of this novel protein was evaluated in a variety of human and non-human primate tissues and was found to be highly restricted. Expression was present only in alveolar macrophages in the lung and in Kupffer cells of the hepatic sinusoids. Expression in these cells was significantly increased when these distinct cell populations were activated. Though these two subpopulations of tissue macrophages are located in different organs, they have similar biological functions. Both types of these phagocytes act as biological filters to remove material from the blood stream or airways including pathogens, senescent cells and proteins and both are capable of secreting a wide variery of important proinflammatory cytokines. In inflamed lung (7 patient samples) expression was prominent in reactive alveolar macrophage cell populations defined as large, pale often vacuolated cells present singly or in aggregates within alveoli and was weak to negative in normal, non-reactive macrophages (single scattered cells of normal size). Expression in alveolar macrophages was increased during inflammation when these cells were both increased in numbers and size (activated). Despite the presence of histocytes in areas of interstial inflammtion and peribronchial lymphoid hyperplasia in these tissues, expression was restricted to alveolar macrophages. Many of the inflamed lungs also had some degree of suppurative inflammation; expression was not present in neutrophilic granulocytes. In liver, there was strong expression in reactive/activated Kupffer cells in livers with acute centrilobular necrosis (acetominophen toxicity) or fairly marked periportal inflammtion. However there was weak or no expression in Kupffer cells in normal liver or in liver with only mild inflammation or mild to moderate lobular hyperplasia/hypertrophy. Thus, as in the lung, there was increased expression in acivated/reative cells. There was no expression of this molecule in histiocytes/macropahges present in inflamed bowel, hyperplastic/reactive tonsil or normal lymph node. The lack of expression in these tissues which all contained histiocytic inflammation or resident macrophage populations strongly supports restricted expression to the unique macrophage subset populations defined as alveolar macrophage and hepatic Kupffer cells. Spleen or bone marrow was not available for evaluation. Human tissues evaluated which had no detectable expression included: Inflammatory bowel disease (7 patient samples with moderate to severe disease), tonsil with reactive hyperplasia, peripheral lymphnode, psoriatic skin (2 patient samples with mild to moderate disease), heart, peripheral nerve. Chimp tissues evaluated which had no detectable expression included: tongue, stomach, thymus.

(11) DNA52196-1348 (PRO733)

Generalized low level signal in many tissues and in many cell types. While endothelial cell expression was observed it was not a prominent feature in either fetal, normal or diseased tissues. Human tissues: moderate expression over fetal liver (mainly hepatocytes), lung, skin, adrenal and heart. Fetal spleen, small intestine, brain and eye are negative. Adult normal kidney, bladder epithelium, lung, adrenal, pancreas, skin - all negative. Expression in adult human liver (normal and diseased), renal tubules in end-stage renal disease, adipose tissue, sarcoma, colon, renal cell carcinoma, hepatocellular carcinoma, squamous cell carcinoma. Non human primate tissues: chimp salivary gland, vessels, stomach, tongue, peripheral nerve, thymus, lymph node, thyroid and parathyroid. Rhesus spinal cord negative, cortical and hippocampal neurones positive.

Deposit of Material

The following materials have been deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville, MD, USA (ATCC):

Material	ATCC Dep. No.	Deposit Date
DNA39987-1184	ATCC 209786	April 21, 1998
DNA40625-1189	ATCC 209788	April 21, 1998
DNA23318-1211	ATCC 209787	April 21, 1998
DNA39979-1213	ATCC 209789	April 21, 1998
DNA40594-1233	ATCC 209617	February 5, 1998
DNA45416-1251	ATCC 209620	February 5, 1998
DNA45419-1252	ATCC 209616	February 5, 1998
DNA52594-1270	ATCC 209679	March 17, 1998
DNA45234-1277	ATCC 209654	March 5, 1998
DNA49624-1279	ATCC 209655	March 5, 1998
DNA48309-1280	ATCC 209656	March 5, 1998
DNA46776-1284	ATCC 209721	March 31, 1998
DNA50980-1286	ATCC 209717	March 31, 1998
DNA50913-1287	ATCC 209716	March 31, 1998
DNA50914-1289	ATCC 209722	March 31, 1998
DNA48296-1292	ATCC 209668	March 11, 1998
DNA32284-1307	ATCC 209670	March 11, 1998
DNA36343-1310	ATCC 209718	March 31, 1998
DNA40571-1315	ATCC 209784	April 21, 1998
DNA41386-1316	ATCC 209703	March 26, 1998
DNA44194-1317	ATCC 209808	April 28, 1998
DNA45415-1318	ATCC 209810	April 28, 1998
DNA44189-1322	ATCC 209699	March 26, 1998
DNA48304-1323	ATCC 209811	April 28, 1998
DNA49152-1324	ATCC 209813	April 28, 1998
DNA49646-1327	ATCC 209705	March 26, 1998
DNA49631-1328	ATCC 209806	April 28, 1998
DNA49645-1347	ATCC 209809	April 28, 1998
DNA45493-1349	ATCC 209805	April 28, 1998
DNA48227-1350	ATCC 209812	April 28, 1998
DNA41404-1352	ATCC 209844	May 6, 1998
DNA44196-1353	ATCC 209847	May 6, 1998
DNA52187-1354	ATCC 209845	May 6, 1998
DNA48328-1355	ATCC 209843	May 6, 1998
DNA56352-1358	ATCC 209846	May 6, 1998
DNA53971-1359	ATCC 209750	April 7, 1998
DNA50919-1361	ATCC 209848	May 6, 1998
DNA44179-1362	ATCC 209851	May 6, 1998

DNA54002-1367	ATCC 209754	April 7, 1998
DNA53906-1368	ATCC 209747	April 7, 1998
DNA52185-1370	ATCC 209861	May 14, 1998
DNA53977-1371	ATCC 209862	May 14, 1998
DNA57253-1382	ATCC 209867	May 14, 1998
DNA58847-1383	ATCC 209879	May 20, 1998
DNA58747-1384	ATCC 209868	May 14, 1998
DNA57689-1385	ATCC 209869	May 14, 1998
DNA23330-1390	ATCC 209775	April 14, 1998
DNA26847-1395	ATCC 209772	April 14, 1998
DNA53974-1401	ATCC 209774	April 14, 1998
DNA57039-1402	ATCC 209777	April 14, 1998
DNA57033-1403	ATCC 209905	May 27, 1998
DNA34353-1428	ATCC 209855	May 12, 1998
DNA45417-1432	ATCC 209910	May 27, 1998
DNA39523-1192	ATCC 209424	October 31, 1997
DNA44205-1285	ATCC 209720	March 31, 1998
DNA50911-1288	ATCC 209714	March 31, 1998
DNA48329-1290	ATCC 209785	April 21, 1998
DNA48306-1291	ATCC 209911	May 27, 1998
DNA48336-1309	ATCC 209669	March 11, 1998
DNA44184-1319	ATCC 209704	March 26, 1998
DNA48314-1320	ATCC 209702	March 26, 1998
DNA48333-1321	ATCC 209701	March 26, 1998
DNA50920-1325	ATCC 209700	March 26, 1998
DNA50988-1326	ATCC 209814	April 28, 1998
DNA48331-1329	ATCC 209715	March 31, 1998
DNA30867-1335	ATCC 209807	April 28, 1998
DNA55737-1345	ATCC 209753	April 7, 1998
DNA49829-1346	ATCC 209749	April 7, 1998
DNA52196-1348	ATCC 209748	April 7, 1998
DNA56965-1356	ATCC 209842	May 6, 1998
DNA56405-1357	ATCC 209849	May 6, 1998
DNA57530-1375	ATCC 209880	May 20, 1998
DNA56439-1376	ATCC 209864	May 14, 1998
DNA56409-1377	ATCC 209882	May 20, 1998
DNA56112-1379	ATCC 209883	May 20, 1998
DNA56045-1380	ATCC 209865	May 14, 1998
DNA59294-1381	ATCC 209866	May 14, 1998
DNA56433-1406	ATCC 209857	May 12, 1998
DNA53912-1457	ATCC 209870	May 14, 1998
DNA50921-1458	ATCC 209859	May 12, 1998
DNA29101-1122	ATCC 209653	March 5, 1998
DNA40021-1154	ATCC 209389	October 17, 1997
DNA42663-1154	ATCC 209386	October 17, 1997
DNA30943-1-1163-1	ATCC 209791	April 21, 1998
DNA64907-1163-1	ATCC 203242	September 9, 1998
DNA64908-1163-1	ATCC 203243	September 9, 1998
DNA39975-1210	ATCC 209783	April 21, 1998
DNA43316-1237	ATCC 209487	November 21, 1997
DNA55800-1263	ATCC 209680	March 17, 1998

These deposit were made under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure and the Regulations thereunder (Budapest Treaty). This assures maintenance of a viable culture of the deposit for 30 years from the date of deposit. The deposits will be made available by ATCC under the terms of the Budapest Treaty, and subject to an agreement between Genentech, Inc. and ATCC, which assures permanent and unrestricted availability of the progeny of the

culture of the deposit to the public upon issuance of the pertinent U.S. patent or upon laying open to the public of any U.S. or foreign patent application, whichever comes first, and assures availability of the progeny to one determined by the U.S. Commissioner of Patents and Trademarks to be entitled thereto according to 35 USC § 122 and the Commissioner's rules pursuant thereto (including 37 CFR § 1.14 with particular reference to 886 OG 638).

The assignee of the present application has agreed that if a culture of the materials on deposit should die or be lost or destroyed when cultivated under suitable conditions, the materials will be promptly replaced on notification with another of the same. Availability of the deposited material is not to be construed as a license to practice the invention in contravention of the rights granted under the authority of any government in accordance with its patent laws.

The foregoing written specification is considered to be sufficient to enable one skilled in the art to practice the invention. The present invention is not to be limited in scope by the construct deposited, since the deposited embodiment is intended as a single illustration of certain aspects of the invention and any constructs that are functionally equivalent are within the scope of this invention. The deposit of material herein does not constitute an admission that the written description herein contained is inadequate to enable the practice of any aspect of the invention, including the best mode thereof, nor is it to be construed as limiting the scope of the claims to the specific illustrations that it represents. Indeed, various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and fall within the scope of the appended claims.

WHAT IS CLAIMED IS:

1. Isolated nucleic acid having at least 80% sequence identity to a nucleotide sequence that encodes a polypeptide comprising an amino acid sequence selected from the group consisting of the amino acid sequence shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:7), Figure 9 (SEQ ID NO:19), Figure 11 (SEQ ID NO:28), Figure 15 (SEQ ID NO:36), Figure 20 (SEQ ID NO:45), Figure 22 (SEQ ID NO:52), Figure 24 (SEQ ID NO:59), Figure 26 (SEQ ID NO:64), Figure 28 (SEQ ID NO:69), Figure 30 (SEQ ID NO:74), Figure 33 (SEQ ID NO:85), Figure 35 (SEQ ID NO:90), Figure 37 (SEQ ID NO:97), Figure 39 (SEQ ID NO:102), Figure 41 (SEQ ID NO:109), Figure 43 (SEQ ID NO:114), Figure 45 (SEQ ID NO:119), Figure 47 (SEQ ID NO:124), Figure 49 (SEQ ID NO:132), Figure 51 (SEQ ID NO:137), Figure 53 (SEQ ID NO:145), Figure 55 (SEQ ID NO:150), Figure 59 (SEQ ID NO:157), Figure 61 (SEQ ID NO:162), Figure 63 (SEQ ID NO:169), Figure 66 (SEQ ID NO:178), Figure 68 (SEQ ID NO:183), Figure 70 (SEQ ID NO:190), Figure 73 (SEQ ID NO:196), Figure 75 (SEQ ID NO:206), Figure 77 (SEQ ID NO:211), Figure 79 (SEQ ID NO:216), Figure 81 (SEQ ID NO:221), Figure 83 (SEQ ID NO:226), Figure 85 (SEQ ID NO:231), Figure 87 (SEQ ID NO:236), Figure 89 (SEQ ID NO:245), Figure 91 (SEQ ID NO:254), Figure 93 (SEQ ID NO:259), Figure 95 (SEQ ID NO:264), Figure 98 (SEQ ID NO:270), Figure 109 (SEQ ID NO:284), Figure 118 (SEQ ID NO:296), Figure 120 (SEQ ID NO:301), Figure 122 (SEQ ID NO:303), Figure 125 (SEQ ID NO:309), Figure 129 (SEQ ID NO:322), Figure 132 (SEQ ID NO:330), Figure 136 (SEQ ID NO:337), Figure 139 (SEQ ID NO:346), Figure 142 (SEQ ID NO:352), Figure 145 (SEQ ID NO:358), Figure 147 (SEQ ID NO:363), Figure 149 (SEQ ID NO:370), Figure 151 (SEQ ID NO:375), Figure 153 (SEQ ID NO:380), Figure 155 (SEQ ID NO:385), Figure 157 (SEQ ID NO:390), Figure 159 (SEQ ID NO:395), Figure 161 (SEQ ID NO:400), Figure 163 (SEQ ID NO:405), Figure 165 (SEQ ID NO:410), Figure 167 (SEQ ID NO:415), Figure 169 (SEQ ID NO:420), Figure 171 (SEQ ID NO:425), Figure 173 (SEQ ID NO:430), Figure 177 (SEQ ID NO:437), Figure 179 (SEQ ID NO:442), Figure 181 (SEQ ID NO:447), Figure 183 (SEQ ID NO:452), Figure 185 (SEQ ID NO:454), Figure 187 (SEQ ID NO:456), Figure 190 (SEQ ID NO:459), Figure 192 (SEQ ID NO:464), Figure 194 (SEQ ID NO:466), Figure 196 (SEQ ID NO:468), Figure 198 (SEQ ID NO:470), Figure 200 (SEQ ID NO:472), Figure 202 (SEQ ID NO:477), Figure 204 (SEQ ID NO:483), Figure 207 (SEQ ID NO:488), Figure 209 (SEQ ID NO:496), Figure 211 (SEQ ID NO:498), Figure 213 (SEQ ID NO:506), Figure 215 (SEQ ID NO:508), Figure 217 (SEQ ID NO:510), Figure 219 (SEQ ID NO:515), Figure 222 (SEQ ID NO:523) and Figure 225 (SEQ ID NO:526).

2. The nucleic acid sequence of Claim 1, wherein said nucleotide sequence comprises a nucleotide sequence selected from the group consisting of the sequence shown in Figure 1 (SEQ ID NO:1), Figure 3 (SEQ ID NO:6), Figure 8 (SEQ ID NO:18), Figure 10 (SEQ ID NO:27), Figure 14 (SEQ ID NO:35), Figure 19 (SEQ ID NO:44), Figure 21 (SEQ ID NO:51), Figure 23 (SEQ ID NO:58), Figure 25 (SEQ ID NO:63), Figure 27 (SEQ ID NO:68), Figure 29 (SEQ ID NO:73), Figure 32 (SEQ ID NO:84), Figure 34 (SEQ ID NO:89), Figure 36 (SEQ ID NO:96), Figure 38 (SEQ ID NO:101), Figure 40 (SEQ ID NO:108), Figure 42 (SEQ ID NO:113), Figure 44 (SEQ ID NO:118), Figure 46 (SEQ ID NO:123), Figure 48 (SEQ ID NO:131), Figure 50 (SEQ ID NO:136), Figure 52 (SEQ ID NO:144), Figure 54 (SEQ ID NO:149), Figure 58 (SEQ ID NO:156), Figure 60 (SEQ ID NO:161), Figure 62 (SEQ ID NO:168), Figure 65 (SEQ ID NO:177), Figure 67 (SEQ ID NO:182), Figure 69 (SEQ ID NO:189), Figure 72 (SEQ ID NO:195), Figure 74 (SEQ ID NO:205), Figure 76 (SEQ ID NO:210), Figure 78 (SEQ ID NO:215), Figure 80 (SEQ ID NO:220), Figure 82 (SEQ ID NO:225), Figure 84 (SEQ ID NO:230), Figure 86 (SEQ ID NO:235), Figure 88 (SEQ ID NO:244), Figure 90 (SEQ ID NO:253), Figure 92 (SEQ ID NO:258), Figure 94

(SEQ ID NO:263), Figure 97 (SEQ ID NO:269), Figure 108 (SEQ ID NO:283), Figure 117 (SEQ ID NO:295), Figure 119 (SEQ ID NO:300), Figure 121 (SEQ ID NO:302), Figure 124 (SEQ ID NO:308), Figure 128 (SEQ ID NO:321), Figure 131 (SEQ ID NO:329), Figure 135 (SEQ ID NO:336), Figure 138 (SEQ ID NO:345), Figure 141 (SEQ ID NO:351), Figure 144 (SEQ ID NO:357), Figure 146 (SEQ ID NO:362), Figure 148 (SEQ ID NO:369), Figure 150 (SEQ ID NO:374), Figure 152 (SEQ ID NO:379), Figure 154 (SEQ ID NO:384), Figure 156 (SEQ ID NO:389), Figure 158 (SEQ ID NO:394), Figure 160 (SEQ ID NO:399), Figure 162 (SEQ ID NO:404), Figure 164 (SEQ ID NO:409), Figure 166 (SEQ ID NO:414), Figure 168 (SEQ ID NO:419), Figure 170 (SEQ ID NO:424), Figure 172 (SEQ ID NO:429), Figure 176 (SEQ ID NO:436), Figure 178 (SEQ ID NO:441), Figure 180 (SEQ ID NO:446), Figure 182 (SEQ ID NO:451), Figure 184 (SEQ ID NO:453), Figure 186 (SEQ ID NO:455), Figure 189 (SEQ ID NO:458), Figure 191 (SEQ ID NO:463), Figure 193 (SEQ ID NO:465), Figure 195 (SEQ ID NO:467), Figure 197 (SEQ ID NO:469), Figure 199 (SEQ ID NO:471), Figure 201 (SEQ ID NO:476), Figure 203 (SEQ ID NO:482), Figure 206 (SEQ ID NO:487), Figure 208 (SEQ ID NO:495), Figure 210 (SEQ ID NO:497), Figure 212 (SEQ ID NO:505), Figure 214 (SEQ ID NO:507), Figure 216 (SEQ ID NO:509), Figure 218 (SEQ ID NO:514), Figure 221 (SEQ ID NO:522) and Figure 224 (SEQ ID NO:525).

3. The nucleic acid of Claim 1, wherein said nucleotide sequence comprises a nucleotide sequence selected from the group consisting of the full-length coding sequence of the sequence shown in Figure 1 (SEQ ID NO:1), Figure 3 (SEQ ID NO:6), Figure 8 (SEQ ID NO:18), Figure 10 (SEQ ID NO:27), Figure 14 (SEQ ID NO:35), Figure 19 (SEQ ID NO:44), Figure 21 (SEQ ID NO:51), Figure 23 (SEQ ID NO:58), Figure 25 (SEQ ID NO:63), Figure 27 (SEQ ID NO:68), Figure 29 (SEQ ID NO:73), Figure 32 (SEQ ID NO:84), Figure 34 (SEQ ID NO:89), Figure 36 (SEQ ID NO:96), Figure 38 (SEQ ID NO:101), Figure 40 (SEQ ID NO:108), Figure 42 (SEQ ID NO:113), Figure 44 (SEQ ID NO:118), Figure 46 (SEQ ID NO:123), Figure 48 (SEQ ID NO:131), Figure 50 (SEQ ID NO:136), Figure 52 (SEQ ID NO:144), Figure 54 (SEQ ID NO:149), Figure 58 (SEQ ID NO:156), Figure 60 (SEQ ID NO:161), Figure 62 (SEQ ID NO:168), Figure 65 (SEQ ID NO:177), Figure 67 (SEQ ID NO:182), Figure 69 (SEQ ID NO:189), Figure 72 (SEQ ID NO:195), Figure 74 (SEQ ID NO:205), Figure 76 (SEQ ID NO:210), Figure 78 (SEQ ID NO:215), Figure 80 (SEQ ID NO:220), Figure 82 (SEQ ID NO:225), Figure 84 (SEQ ID NO:230), Figure 86 (SEQ ID NO:235), Figure 88 (SEQ ID NO:244), Figure 90 (SEQ ID NO:253), Figure 92 (SEQ ID NO:258), Figure 94 (SEQ ID NO:263), Figure 97 (SEQ ID NO:269), Figure 108 (SEQ ID NO:283), Figure 117 (SEQ ID NO:295), Figure 119 (SEQ ID NO:300), Figure 121 (SEQ ID NO:302), Figure 124 (SEQ ID NO:308), Figure 128 (SEQ ID NO:321), Figure 131 (SEQ ID NO:329), Figure 135 (SEQ ID NO:336), Figure 138 (SEQ ID NO:345), Figure 141 (SEQ ID NO:351), Figure 144 (SEQ ID NO:357), Figure 146 (SEQ ID NO:362), Figure 148 (SEQ ID NO:369), Figure 150 (SEQ ID NO:374), Figure 152 (SEQ ID NO:379), Figure 154 (SEQ ID NO:384), Figure 156 (SEQ ID NO:389), Figure 158 (SEQ ID NO:394), Figure 160 (SEQ ID NO:399), Figure 162 (SEQ ID NO:404), Figure 164 (SEQ ID NO:409), Figure 166 (SEQ ID NO:414), Figure 168 (SEQ ID NO:419), Figure 170 (SEQ ID NO:424), Figure 172 (SEQ ID NO:429), Figure 176 (SEQ ID NO:436), Figure 178 (SEQ ID NO:441), Figure 180 (SEQ ID NO:446), Figure 182 (SEQ ID NO:451), Figure 184 (SEQ ID NO:453), Figure 186 (SEQ ID NO:455), Figure 189 (SEQ ID NO:458), Figure 191 (SEQ ID NO:463), Figure 193 (SEQ ID NO:465), Figure 195 (SEQ ID NO:467), Figure 197 (SEQ ID NO:469), Figure 199 (SEQ ID NO:471), Figure 201 (SEQ ID NO:476), Figure 203 (SEQ ID NO:482), Figure 206 (SEQ ID NO:487), Figure 208 (SEQ ID NO:495), Figure 210 (SEQ ID NO:497), Figure 212 (SEQ ID NO:505), Figure 214 (SEQ ID NO:507), Figure 216 (SEQ ID NO:509), Figure 218

(SEQ ID NO:514), Figure 221 (SEQ ID NO:522) or Figure 224 (SEQ ID NO:525).

4. Isolated nucleic acid which comprises the full-length coding sequence of the DNA deposited under accession number ATCC 209791, ATCC 209786, ATCC 209788, ATCC 209787, ATCC 209789, ATCC 209617, ATCC 209620, ATCC 209616, ATCC 209679, ATCC 209654, ATCC 209655, ATCC 209656, ATCC 209721, ATCC 209717, ATCC 209716, ATCC 209722, ATCC 209668, ATCC 209670, ATCC 209718, ATCC 209784, ATCC 209703, ATCC 209808, ATCC 209810, ATCC 209699, ATCC 209811, ATCC 209813, ATCC 209705, ATCC 209806, ATCC 209809, ATCC 209805, ATCC 209812, ATCC 209844, ATCC 209847, ATCC 209845, ATCC 209843, ATCC 209846, ATCC 209750, ATCC 209848, ATCC 209851, ATCC 209754, ATCC 209747, ATCC 209861, ATCC 209862, ATCC 209867, ATCC 209879, ATCC 209868, ATCC 209869, ATCC 209775, ATCC 209772, ATCC 209774, ATCC 209777, ATCC 209905, ATCC 209855, ATCC 209910, ATCC 209424, ATCC 209720, ATCC 209714, ATCC 209785, ATCC 209911, ATCC 209669, ATCC 209704, ATCC 209702, ATCC 209701, ATCC 209700, ATCC 209814, ATCC 209715, ATCC 209807, ATCC 209753, ATCC 209749, ATCC 209748, ATCC 209842, ATCC 209849, ATCC 209880, ATCC 209864, ATCC 209882, ATCC 209883, ATCC 209865, ATCC 209866, ATCC 209857, ATCC 209870, ATCC 209859, ATCC 209653, ATCC 209389, ATCC 209386, ATCC 203242, ATCC 203243, ATCC 209783, ATCC 209487 or ATCC 209680.

5. A vector comprising the nucleic acid of Claim 1.

6. The vector of Claim 5 operably linked to control sequences recognized by a host cell transformed with the vector.

7. A host cell comprising the vector of Claim 5.

8. The host cell of Claim 7 wherein said cell is a CHO cell.

9. The host cell of Claim 7 wherein said cell is an *E. coli*.

10. The host cell of Claim 7 wherein said cell is a yeast cell.

11. A process for producing a PRO polypeptides comprising culturing the host cell of Claim 7 under conditions suitable for expression of said PRO polypeptide and recovering said PRO polypeptide from the cell culture.

12. Isolated native sequence PRO polypeptide having at least 80% sequence identity to an amino acid sequence selected from the group consisting of the amino acid sequence shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:7), Figure 9 (SEQ ID NO:19), Figure 11 (SEQ ID NO:28), Figure 15 (SEQ ID NO:36), Figure 20 (SEQ ID NO:45), Figure 22 (SEQ ID NO:52), Figure 24 (SEQ ID NO:59), Figure 26 (SEQ ID NO:64), Figure 28 (SEQ ID NO:69), Figure 30 (SEQ ID NO:74), Figure 33 (SEQ ID NO:85), Figure 35 (SEQ ID NO:90), Figure 37 (SEQ ID NO:97), Figure 39 (SEQ ID NO:102), Figure 41 (SEQ ID NO:109), Figure 43 (SEQ ID NO:114), Figure 45 (SEQ ID NO:119), Figure 47 (SEQ ID NO:124), Figure 49 (SEQ ID NO:132), Figure 51 (SEQ ID NO:137),

Figure 53 (SEQ ID NO:145), Figure 55 (SEQ ID NO:150), Figure 59 (SEQ ID NO:157), Figure 61 (SEQ ID NO:162), Figure 63 (SEQ ID NO:169), Figure 66 (SEQ ID NO:178), Figure 68 (SEQ ID NO:183), Figure 70 (SEQ ID NO:190), Figure 73 (SEQ ID NO:196), Figure 75 (SEQ ID NO:206), Figure 77 (SEQ ID NO:211), Figure 79 (SEQ ID NO:216), Figure 81 (SEQ ID NO:221), Figure 83 (SEQ ID NO:226), Figure 85 (SEQ ID NO:231), Figure 87 (SEQ ID NO:236), Figure 89 (SEQ ID NO:245), Figure 91 (SEQ ID NO:254), Figure 93 (SEQ ID NO:259), Figure 95 (SEQ ID NO:264), Figure 98 (SEQ ID NO:270), Figure 109 (SEQ ID NO:284), Figure 118 (SEQ ID NO:296), Figure 120 (SEQ ID NO:301), Figure 122 (SEQ ID NO:303), Figure 125 (SEQ ID NO:309), Figure 129 (SEQ ID NO:322), Figure 132 (SEQ ID NO:330), Figure 136 (SEQ ID NO:337), Figure 139 (SEQ ID NO:346), Figure 142 (SEQ ID NO:352), Figure 145 (SEQ ID NO:358), Figure 147 (SEQ ID NO:363), Figure 149 (SEQ ID NO:370), Figure 151 (SEQ ID NO:375), Figure 153 (SEQ ID NO:380), Figure 155 (SEQ ID NO:385), Figure 157 (SEQ ID NO:390), Figure 159 (SEQ ID NO:395), Figure 161 (SEQ ID NO:400), Figure 163 (SEQ ID NO:405), Figure 165 (SEQ ID NO:410), Figure 167 (SEQ ID NO:415), Figure 169 (SEQ ID NO:420), Figure 171 (SEQ ID NO:425), Figure 173 (SEQ ID NO:430), Figure 177 (SEQ ID NO:437), Figure 179 (SEQ ID NO:442), Figure 181 (SEQ ID NO:447), Figure 183 (SEQ ID NO:452), Figure 185 (SEQ ID NO:454), Figure 187 (SEQ ID NO:456), Figure 190 (SEQ ID NO:459), Figure 192 (SEQ ID NO:464), Figure 194 (SEQ ID NO:466), Figure 196 (SEQ ID NO:468), Figure 198 (SEQ ID NO:470), Figure 200 (SEQ ID NO:472), Figure 202 (SEQ ID NO:477), Figure 204 (SEQ ID NO:483), Figure 207 (SEQ ID NO:488), Figure 209 (SEQ ID NO:496), Figure 211 (SEQ ID NO:498), Figure 213 (SEQ ID NO:506), Figure 215 (SEQ ID NO:508), Figure 217 (SEQ ID NO:510), Figure 219 (SEQ ID NO:515), Figure 222 (SEQ ID NO:523) and Figure 225 (SEQ ID NO:526).

13. Isolated PRO polypeptide having at least 80% sequence identity to the amino acid sequence encoded by the nucleotide deposited under accession number ATCC 209791, ATCC 209786, ATCC 209788, ATCC 209787, ATCC 209789, ATCC 209617, ATCC 209620, ATCC 209616, ATCC 209679, ATCC 209654, ATCC 209655, ATCC 209656, ATCC 209721, ATCC 209717, ATCC 209716, ATCC 209722, ATCC 209668, ATCC 209670, ATCC 209718, ATCC 209784, ATCC 209703, ATCC 209808, ATCC 209810, ATCC 209699, ATCC 209811, ATCC 209813, ATCC 209705, ATCC 209806, ATCC 209809, ATCC 209805, ATCC 209812, ATCC 209844, ATCC 209847, ATCC 209845, ATCC 209843, ATCC 209846, ATCC 209750, ATCC 209848, ATCC 209851, ATCC 209754, ATCC 209747, ATCC 209861, ATCC 209862, ATCC 209867, ATCC 209879, ATCC 209868, ATCC 209869, ATCC 209775, ATCC 209772, ATCC 209774, ATCC 209777, ATCC 209905, ATCC 209855, ATCC 209910, ATCC 209424, ATCC 209720, ATCC 209714, ATCC 209785, ATCC 209911, ATCC 209669, ATCC 209704, ATCC 209702, ATCC 209701, ATCC 209700, ATCC 209814, ATCC 209715, ATCC 209807, ATCC 209753, ATCC 209749, ATCC 209748, ATCC 209842, ATCC 209849, ATCC 209880, ATCC 209864, ATCC 209882, ATCC 209883, ATCC 209865, ATCC 209866, ATCC 209857, ATCC 209870, ATCC 209859, ATCC 209653, ATCC 209389, ATCC 209386, ATCC 203242, ATCC 203243, ATCC 209783, ATCC 209487 or ATCC 209680.

14. A chimeric molecule comprising a polypeptide according to Claim 12 fused to a heterologous amino acid sequence.

15. The chimeric molecule of Claim 14 wherein said heterologous amino acid sequence is an epitope

tag sequence.

16. The chimeric molecule of Claim 14 wherein said heterologous amino acid sequence is a Fc region of an immunoglobulin.

17. An antibody which specifically binds to a PRO polypeptide according to Claim 12.

18. The antibody of Claim 17 wherein said antibody is a monoclonal antibody.

19. An isolated nucleic acid molecule which has at least 80% sequence identity to a nucleic acid which comprises a nucleotide sequence selected from the group consisting of that shown in Figure 5 (SEQ ID NO:8), Figure 6 (SEQ ID NO:9), Figure 7 (SEQ ID NO:10), Figure 12 (SEQ ID NO:29), Figure 13 (SEQ ID NO:30), Figure 16 (SEQ ID NO:37), Figure 17 (SEQ ID NO:38), Figure 18 (SEQ ID NO:39), Figure 31 (SEQ ID NO:75), Figure 64 (SEQ ID NO:170), Figure 71 (SEQ ID NO:191), Figure 96 (SEQ ID NO:265), Figure 99 (SEQ ID NO:271), Figure 100 (SEQ ID NO:272), Figure 101 (SEQ ID NO:273), Figure 102 (SEQ ID NO:274), Figure 103 (SEQ ID NO:275), Figure 104 (SEQ ID NO:276), Figure 105 (SEQ ID NO:277), Figure 106 (SEQ ID NO:278), Figure 107 (SEQ ID NO:279), Figure 110 (SEQ ID NO:285), Figure 111 (SEQ ID NO:286), Figure 112 (SEQ ID NO:287), Figure 113 (SEQ ID NO:288), Figure 114 (SEQ ID NO:289), Figure 115 (SEQ ID NO:290), Figure 116 (SEQ ID NO:291), Figure 123 (SEQ ID NO:304), Figure 126 (SEQ ID NO:310), Figure 127 (SEQ ID NO:311), Figure 130 (SEQ ID NO:323), Figure 133 (SEQ ID NO:331), Figure 134 (SEQ ID NO:332), Figure 137 (SEQ ID NO:338), Figure 140 (SEQ ID NO:347), Figure 143 (SEQ ID NO:353), Figure 174 (SEQ ID NO:431), Figure 175 (SEQ ID NO:432), Figure 188 (SEQ ID NO:457), Figure 205 (SEQ ID NO:484), Figure 220 (SEQ ID NO:516), Figure 223 (SEQ ID NO:524), Figure 226 (SEQ ID NO:527), Figure 227 (SEQ ID NO:528) and Figure 228 (SEQ ID NO:529).

FIGURE 1

CCAGGTCCAACTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCT
CGACCTCGACCCACGCGTCCGCCAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGA
CAGGCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGC
AAGGGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGC
TCCAGCAGCATCAGCAGCCCCCAGGACCGGGGGAGGCACAGGTGGCCCCCACCACCCGGAGG
AGCAGCTCCTGCCCCTGTCCGGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGA
AGGCCACCCCGCCTGGAGGCACAGGCCATGAGGGGCTCTCAGGAGGTGCTGCTGATGTGGCT
TCTGGTGTTGGCAGTGGGCGGCACAGAGCACGCCTACCGGCCCGGCCGTTAGGGTGTGTGCT
GTCCCGGGCTCACGGGGACCCTGTCTCCGAGTCGTTCGTGCAGCGTGTGTACCAGCCCTTCC
TCACCACCTGCGACGGGCACCGGGCCTGCAGCACCTACCGAACCATTTATAGGACCGCCTAC
CGCCGCAGCCCTGGGCTGGCCCCTGCCAGGCCTCGCTACGCGTGCTGCCCCGGCTGGAAGAG
GACCAGCGGGCTTCCTGGGGCCTGTGGAGCAGCAATATGCCAGCCGCCATGCCGGAACGGAG
GGAGCTGTGTCCAGCCTGGCCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAG
TCAGATGTGGATGAATGCAGTGCTAGGAGGGGCGGCTGTCCCCAGCGCTGCATCAACACCGC
CGGCAGTTACTGGTGCCAGTGTTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTG
TGCCCAAGGGAGGGCCCCCCAGGGTGGCCCCCAACCCGACAGGAGTGGACAGTGCAATGAAG
GAAGAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGCT
GGCCCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGCATGGGCTCCCGGACCCCGGCAGCC
TCCTGGTGCACTCCTTCCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTTCCTTC
CTGGAGGAGCAGCTGGGGTCCTGCTCCTGCAAGAAAGACTCGTGACTGCCCAGCGCCCCAGG
CTGGACTGAGCCCCTCACGCCGCCCTGCAGCCCCCATGCCCCTGCCCAACATGCTGGGGGTC
CAGAAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCCTCCTTTTCCTCCTC
CCCTTCCCTCGGGAGGGTCCCCAGACCCTGGCATGGGATGGGCTGGGATTTTTTTTGTGAAT
CCACCCCTGGCTACCCCCACCCTGGTTACCCCAACGGCATCCCAAGGCCAGGTGGGCCCTCA
GCTGAGGGAAGGTACGAGTTCCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCC
CGGAGGCTGGGTGGGGCCTCAGTGGGGGCTGCTGCCTGACCCCCAGCACAATAAAAATGAAA
CGTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCT
AGAGTCGACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATGGT
TACAAAT

FIGURE 2

MTDSPPPGHPEEKATPPGGTGHEGLSGGAADVASGVGSGRHRARLPARPLGCVLSRAHGDPV
SESFVQRVYQPFLTTCDGHRACSTYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGAC
GAAICQPPCRNGGSCVQPGRCRCPAGWRGDTCQSDVDECSARRGGCPQRCINTAGSYWCQCW
EGHSLSADGTLCVPKGGPPRVAPNPTGVDSAMKEEVQRLQSRVDLLEEKLQLVLAPLHSLAS
QALEHGLPDPGSLLVHSFQQLGRIDSLSEQISFLEEQLGSCSCKKDS

FIGURE 3

CGCTCGCCCCGTCGCCCCTCGCCTCCCCGCAGAGTCCCCTCGCGGCAGCAGATGTGTGTGGG
GTCAGCCCACGGCGGGGACTATGGTGAAATTCCCGGCGCTCACGCACTACTGGCCCCTGATC
CGGTTCTTGGTGCCCCTGGGCATCACCAACATAGCCATCGACTTCGGGGAGCAGGCCTTGAA
CCGGGGCATTGCTGCTGTCAAGGAGGATGCAGTCGAGATGCTGGCCAGCTACGGGCTGGCGT
ACTCCCTCATGAAGTTCTTCACGGGTCCCATGAGTGACTTCAAAAATGTGGGCCTGGTGTTT
GTGAACAGCAAGAGAGACAGGACCAAAGCCGTCCTGTGTATGGTGGTGGCAGGGGCCATCGC
TGCCGTCTTTCACACACTGATAGCTTATAGTGATTTAGGATACTACATTATCAATAAACTGC
ACCATGTGGACGAGTCGGTGGGGAGCAAGACGAGAAGGGCCTTCCTGTACCTCGCCGCCTTT
CCTTTCATGGACGCAATGGCATGGACCCATGCTGGCATTCTCTTAAAACACAAATACAGTTT
CCTGGTGGGATGTGCCTCAATCTCAGATGTCATAGCTCAGGTTGTTTTTGTAGCCATTTTGC
TTCACAGTCACCTGGAATGCCGGGAGCCCCTGCTCATCCCGATCCTCTCCTTGTACATGGGC
GCACTTGTGCGCTGCACCACCCTGTGCCTGGGCTACTACAAGAACATTCACGACATCATCCC
TGACAGAAGTGGCCCGGAGCTGGGGGGAGATGCAACAATAAGAAAGATGCTGAGCTTCTGGT
GGCCTTTGGCTCTAATTCTGGCCACACAGAGAATCAGTCGGCCTATTGTCAACCTCTTTGTT
TCCCGGGACCTTGGTGGCAGTTCTGCAGCCACAGAGGCAGTGGCGATTTTGACAGCCACATA
CCCTGTGGGTCACATGCCATACGGCTGGTTGACGGAAATCCGTGCTGTGTATCCTGCTTTCG
ACAAGAATAACCCCAGCAACAAACTGGTGAGCACGAGCAACACAGTCACGGCAGCCCACATC
AAGAAGTTCACCTTCGTCTGCATGGCTCTGTCACTCACGCTCTGTTTCGTGATGTTTTGGAC
ACCCAACGTGTCTGAGAAAATCTTGATAGACATCATCGGAGTGGACTTTGCCTTTGCAGAAC
TCTGTGTTGTTCCTTTGCGGATCTTCTCCTTCTTCCCAGTTCCAGTCACAGTGAGGGCGCAT
CTCACCGGGTGGCTGATGACACTGAAGAAAACCTTCGTCCTTGCCCCCAGCTCTGTGCTGCG
GATCATCGTCCTCATCGCCAGCCTCGTGGTCCTACCCTACCTGGGGGTGCACGGTGCGACCC
TGGGCGTGGGCTCCCTCCTGGCGGGCTTTGTGGGAGAATCCACCATGGTCGCCATCGCTGCG
TGCTATGTCTACCGGAAGCAGAAAAAGAAGATGGAGAATGAGTCGGCCACGGAGGGGGAAGA
CTCTGCCATGACAGACATGCCTCCGACAGAGGAGGTGACAGACATCGTGGAAATGAGAGAGG
AGAATGAATAAGGCACGGGACGCCATGGGCACTGCAGGGACGGTCAGTCAGGATGACACTTC
GGCATCATCTCTTCCCTCTCCCATCGTATTTTGTTCCCTTTTTTTTGTTTTGTTTTGGTAAT
GAAAGAGGCCTTGATTTAAAGGTTTCGTGTCAATTCTCTAGCATACTGGGTATGCTCACACT
GACGGGGGGACCTAGTGAATGGTCTTTACTGTTGCTATGTAAAAACAAACGAAACAACTGAC
TTCATACCCCTGCCTCACGAAAACCCAAAAGACACAGCTGCCTCACGGTTGACGTTGTGTCC
TCCTCCCCTGGACAATCTCCTCTTGGAACCAAAGGACTGCAGCTGTGCCATCGCGCCTCGGT
CACCCTGCACAGCAGGCCACAGACTCTCCTGTCCCCCTTCATCGCTCTTAAGAATCAACAGG
TTAAAACTCGGCTTCCTTTGATTTGCTTCCCAGTCACATGGCCGTACAAAGAGATGGAGCCC
CGGTGGCCTCTTAAATTTCCCTTCTGCCACGGAGTTCGAAACCATCTACTCCACACATGCAG
GAGGCGGGTGGCACGCTGCAGCCCGGAGTCCCCGTTCACACTGAGGAACGGAGACCTGTGAC
CACAGCAGGCTGACAGATGGACAGAATCTCCCGTAGAAAGGTTTGGTTTGAAATGCCCCGGG
GGCAGCAAACTGACATGGTTGAATGATAGCATTTCACTCTGCGTTCTCCTAGATCTGAGCAA
GCTGTCAGTTCTCACCCCCACCGTGTATATACATGAGCTAACTTTTTTAAATTGTCACAAAA
GCGCATCTCCAGATTCCAGACCCTGCCGCATGACTTTTCCTGAAGGCTTGCTTTTCCCTCGC
CTTTCCTGAAGGTCGCATTAGAGCGAGTCACATGGAGCATCCTAACTTTGCATTTTAGTTTT
TACAGTGAACTGAAGCTTTAAGTCTCATCCAGCATTCTAATGCCAGGTTGCTGTAGGGTAAC
TTTTGAAGTAGATATATTACCTGGTTCTGCTATCCTTAGTCATAACTCTGCGGTACAGGTAA
TTGAGAATGTACTACGGTACTTCCCTCCCACACCATACGATAAAGCAAGACATTTTATAACG
ATACCAGAGTCACTATGTGGTCCTCCCTGAAATAACGCATTCGAAATCCATGCAGTGCAGTA
TATTTTTCTAAGTTTTGGAAAGCAGGTTTTTTCCTTTAAAAAAATTATAGACACGGTTCACT
AAATTGATTTAGTCAGAATTCCTAGACTGAAAGAACCTAAACAAAAAAATATTTTAAAGATA
TAAATATATGCTGTATATGTTATGTAATTTATTTTAGGCTATAATACATTTCCTATTTTCGC
ATTTTCAATAAAATGTCTCTAATACAAAAAA

FIGURE 4

MVKFPALTHYWPLIRFLVPLGITNIAIDFGEQALNRGIAAVKEDAVEMLASYGLAYSLMKFF
TGPMSDFKNVGLVFVNSKRDRTKAVLCMVVAGAIAAVFHTLIAYSDLGYYIINKLHHVDESV
GSKTRRAFLYLAAFPFMDAMAWTHAGILLKHKYSFLVGCASISDVIAQVVFVAILLHSHLEC
REPLLIPILSLYMGALVRCTTLCLGYYKNIHDIIPDRSGPELGGDATIRKMLSFWWPLALIL
ATQRISRPIVNLFVSRDLGGSSAATEAVAILTATYPVGHMPYGWLTEIRAVYPAFDKNNPSN
KLVSTSNTVTAAHIKKFTFVCMALSLTLCFVMFWTPNVSEKILIDIIGVDFAFAELCVVPLR
IFSFFPVPVTVRAHLTGWLMTLKKTFVLAPSSVLRIIVLIASLVVLPYLGVHGATLGVGSLL
AGFVGESTMVAIAACYVYRKQKKKMENESATEGEDSAMTDMPPTEEVTDIVEMREENE

FIGURE 5

CCTGACAGAAGTGCCCCGGAGCTGGGGGAGATNCAACATTAAGAAGATGCTGAGCTTCTGGT
GCCNTTTGGCTCTAATTCTGGCCACACAGAGAANCAGTCGGCCTATTGTCAACCTCTTTGTT
TCCCGGGACCTTGGTGGCAGTTCTGCAGCCACAGAGGCAGTGGCGATTTTGACAGCCACATA
CCCTGTGGGTCACATGCCATACGGCTGGTTGACGGAAATCCGTGCTGTGTATCCTGCTTTCG
ACAAGAATAACCCCAGCAACAAACTGGTGAGCACGAGCAACACAGTCACGGCGGCCCACATC
AAGAAGTTCACCTTCGTCTGCATGGCTCTGTCACTCACGCTCTGTTTCGTGATGTTTTGGAC
ACCCAACGTGTCTGNGAAAATCTTGATAGACATCATCGGAGTGGACTTTGCCTTTGCAGAAC
TCTGTGTTGTTCCTTTGCGGATCTTCTCCTTCTTCCCAGTTCCAGTCACAGTGAGGGCGCAT
CTCACCGGGTGGCTGATGACACTGAAGAAAACCTTCGTC

FIGURE 6

TGACGGAATCCCGGGCTGGGTATCCTGGTTTNGACAAGATAAACCCCCAGCAANAAATTGGG
GAGCAGGGCAAAACAGTNACGGGCAGCCCACATCAAGAAGTTCACCTTNGTTTGNATGGNTC
TGTCAACTCACGCTNTGTTTCGTGATGTTTTGGACACCCAAAGTGTTTGAGAAAATTTTGAT
AGACATNATCGGAGTGGANTTTGCCTTTGCAGAANTTTGNGNTGTTCCTTTGCGGATTTTCT
CCTTTTTCCCAGTTCCAGTCACAGNGAGGGCGCATCTCACCGGGNGGNTGATGACANTGAAG
AAAACCTTTGTCCTTGCCCCCAGCTNTTTGGTGCGGATCATTGTCCTNATNGCCAGCCTTGT
GGTCCTACCCTACCTGGGGGTGCACGGTGCGACCCTGGGCGTGGGTTCCCTCCTGGCGGGCA

FIGURE 7

TATTCCCAGTTCCGGTCACGGGGAGGGCGCATNTCACCGGGTGGCTGANGACACTGAAGAAA
ACCTTNGTCCTTGCCCCCAGNTTTGTGNTGCGGATNATCGTCCTCATCGCCAGCCTNGTGGT
CCTACCCTACCTGGGGGTGCACGGTGAGAC

FIGURE 8

GCCCCGCGCCCGGCGCCGGGCGCCCGAAGCCGGGAGCCACCGCC**ATG**GGGGCCTGCCTGGGA
GCCTGCTCCCTGCTCAGCTGCGCGTCCTGCCTCTGCGGCTCTGCCCCCTGCATCCTGTGCAG
CTGCTGCCCCGCCAGCCGCAACTCCACCGTGAGCCGCCTCATCTTCACGTTCTTCCTCTTCC
TGGGGGTGCTGGTGTCCATCATTATGCTGAGCCCGGGCGTGGAGAGTCAGCTCTACAAGCTG
CCCTGGGTGTGTGAGGAGGGGGCCGGGATCCCCACCGTCCTGCAGGGCCACATCGACTGTGG
CTCCCTGCTTGGCTACCGCGCTGTCTACCGCATGTGCTTCGCCACGGCGGCCTTCTTCTTCT
TCTTTTTCACCCTGCTCATGCTCTGCGTGAGCAGCAGCCGGGACCCCCGGGCTGCCATCCAG
AATGGGTTTTGGTTCTTTAAGTTCCTGATCCTGGTGGGCCTCACCGTGGGTGCCTTCTACAT
CCCTGACGGCTCCTTCACCAACATCTGGTTCTACTTCGGCGTCGTGGGCTCCTTCCTCTTCA
TCCTCATCCAGCTGGTGCTGCTCATCGACTTTGCGCACTCCTGGAACCAGCGGTGGCTGGGC
AAGGCCGAGGAGTGCGATTCCCGTGCCTGGTACGCAGGCCTCTTCTTCTTCACTCTCCTCTT
CTACTTGCTGTCGATCGCGGCCGTGGCGCTGATGTTCATGTACTACACTGAGCCCAGCGGCT
GCCACGAGGGCAAGGTCTTCATCAGCCTCAACCTCACCTTCTGTGTCTGCGTGTCCATCGCT
GCTGTCCTGCCCAAGGTCCAGGACGCCCAGCCCAACTCGGGTCTGCTGCAGGCCTCGGTCAT
CACCCTCTACACCATGTTTGTCACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCA
ACCCCCATTTGCCAACCCAGCTGGGCAACGAGACAGTTGTGGCAGGCCCCGAGGGCTATGAG
ACCCAGTGGTGGGATGCCCCGAGCATTGTGGGCCTCATCATCTTCCTCCTGTGCACCCTCTT
CATCAGTCTGCGCTCCTCAGACCACCGGCAGGTGAACAGCCTGATGCAGACCGAGGAGTGCC
CACCTATGCTAGACGCCACACAGCAGCAGCAGCAGCAGGTGGCAGCCTGTGAGGGCCGGGCC
TTTGACAACGAGCAGGACGGCGTCACCTACAGCTACTCCTTCTTCCACTTCTGCCTGGTGCT
GGCCTCACTGCACGTCATGATGACGCTCACCAACTGGTACAAGCCCGGTGAGACCCGG**AAGA**
TGATCAGCACGTGGACCGCCGTGTGGGTGAAGATCTGTGCCAGCTGGGCAGGGCTGCTCCTC
TACCTGTGGACCCTGGTAGCCCCACTCCTCCTGCGCAACCGCGACTTCAGC**TGA**GGCAGCCT
CACAGCCTGCCATCTGGTGCCTCCTGCCACCTGGTGCCTCTCGGCTCGGTGACAGCCAACCT
GCCCCCTCCCCACACCAATCAGCCAGGCTGAGCCCCCACCCCTGCCCCAGCTCCAGGACCTG
CCCCTGAGCCGGGCCTTCTAGTCGTAGTGCCTTCAGGGTCCGAGGAGCATCAGGCTCCTGCA
GAGCCCCATCCCCCCGCCACACCCACACGGTGGAGCTGCCTCTTCCTTCCCCTCCTCCCTGT
TGCCCATACTCAGCATCTCGGATGAAAGGGCTCCCTTGTCCTCAGGCTCCACGGGAGCGGGG
CTGCTGGAGAGAGCGGGGAACTCCCACCACAGTGGGGCATCCGGCACTGAAGCCCTGGTGTT
CCTGGTCACGTCCCCCAGGGGACCCTGCCCCCTTCCTGGACTTCGTGCCTTACTGAGTCTCT
AAGACTTTTTCTAATAAACAAGCCAGTGCGTGTAAAAAAAA

FIGURE 9

MGACLGACSLLSCASCLCGSAPCILCSCCPASRNSTVSRLIFTFFLFLGVLVSIIMLSPGVE
SQLYKLPWVCEEGAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFTLLMLCVSSSRD
PRAAIQNGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSFLFILIQLVLLIDFAHSW
NQRWLGKAEECDSRAWYAGLFFFTLLFYLLSIAAVALMFMYYTEPSGCHEGKVFISLNLTFC
VCVSIAAVLPKVQDAQPNSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVA
GPEGYETQWWDAPSIVGLIIFLLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQQVA
ACEGRAFDNEQDGVTYSYSFFHFCLVLASLHVMMTLTNWYKPGETRKMISTWTAVWVKICAS
WAGLLLYLWTLVAPLLLRNRDFS

FIGURE 10

GAGCGAGGCCGGGGACTGAAGGTGTGGGTGTCGAGCCCTCTGGCAGAGGGTTAACCTGGGTC
AAATGCACGGATTCTCACCTCGTACAGTTACGCTCTCCCGCGGCACGTCCGCGAGGACTTGA
AGTCCTGAGCGCTCAAGTTTGTCCGTAGGTCGAGAGAAGGCC**ATG**GAGGTGCCGCCACCGGC
ACCGCGGAGCTTTCTCTGTAGAGCATTGTGCCTATTTCCCCGAGTCTTTGCTGCCGAAGCTG
TGACTGCCGATTCGGAAGTCCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGTCCCAGAGCCC
TATTACCCGGAATCTGGATGGGACCGCCTCCGGGAGCTGTTTGGCAAAGATGAACAGCAGAG
AATTTCAAAGGACCTTGCTAATATCTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGG
TGTATGGGGGAATACCAGCTTTTATTCATGCTAAACAACAATACATTGAGCAGAGCCAGGCA
GAAATTTATCATAACCGGTTTGATGCTGTGCAATCTGCACATCGTGCTGCCACACGAGGCTT
CATTCGTTATGGCTGGCGCTGGGGTTGGAGAACTGCAGTGTTTGTGACTATATTCAACACAG
TGAACACTAGTCTGAATGTATACCGAAATAAAGATGCCTTAAGCCATTTTGTAATTGCAGGA
GCTGTCACGGGAAGTCTTTTTAGGATAAACGTAGGCCTGCGTGGCCTGGTGGCTGGTGGCAT
AATTGGAGCCTTGCTGGGCACTCCTGTAGGAGGCCTGCTGATGGCATTTCAGAAGTACGCTG
GTGAGACTGTTCAGGAAAGAAAACAGAAGGATCGAAAGGCACTCCATGAGCTAAAACTGGAA
GAGTGGAAAGGCAGACTACAAGTTACTGAGCACCTCCCTGAGAAAATTGAAAGTAGTTTACG
GGAAGATGAACCTGAGAATGATGCTAAGAAAATTGAAGCACTGCTAAACCTTCCTAGAAACC
CTTCAGTAATAGATAAACAAGACAAGGAC**TGA**AAGTGCTCTGAACTTGAAACTCACTGGAGA
GCTGAAGGGAGCTGCCATGTCCGATGAATGCCAACAGACAGGCCACTCTTTGGTCAGCCTGC
TGACAAATTTAAGTGCTGGTACCTGTGGTGGCAGTGGCTTGCTCTTGTCTTTTTCTTTTCTT
TTTAACTAAGAATGGGGCTGTTGTACTCTCACTTTACTTATCCTTAAATTTAAATACATACT
TATGTTTGTATTAATCTATCAATATATGCATACATGGATATATCCACCCACCTAGATTTTAA
GCAGTAAATAAAACATTTCGCAAAAGATTAAAGTTGAATTTTACAGTTT

FIGURE 11

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA23318
><subunit 1 of 1, 285 aa, 1 stop
><MW: 32190, pI: 9.03, NX(S/T): 2
MEVPPPAPRSFLCRALCLFPRVFAAEAVTADSEVLEERQKRLPYVPEPYYPESGWDRLRELF
GKDEQQRISKDLANICKTAATAGIIGWVYGGIPAFIHAKQQYIEQSQAEIYHNRFDAVQSAH
RAATRGFIRYGWRWGWRTAVFVTIFNTVNTSLNVYRNKDALSHFVIAGAVTGSLFRINVGLR
GLVAGGIIGALLGTPVGGLLMAFQKYAGETVQERKQKDRKALHELKLEEWKGRLQVTEHLPE
KIESSLREDEPENDAKKIEALLNLPRNPSVIDKQDKD

Important Features:

Signal Peptide:

amino acids 1-24

Transmembrane domains:

amino acids 76-96 and amino acids 171-195

N-glycosylation site:

amino acids 153-156

FIGURE 12

CGGAAGTCCCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGTCCCAGAGCCCTATTACCCGGA
ATCTGGATGGGACCGCTCCGGGAGCTGTTTGGCAAAGATGAACAGCAGAGAATTTCAAAGGA
CCTTGCTAATATCTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGGTGTATGGGGGAA
TACCAGCTTTTATTCATGCTAAACAACAATACATTGAGCAGAGCCAGGCAGAAATTTATCAT
AACCGGTTTGATGCTGTGCAATCTGCACATCGTGCTGCCACACGAGGCTTCATTCGTTCATG
GCTGGCGCCGAACC

FIGURE 13

TCAAGTTTGTCCGTAGGTCGAGAGAAGGCCATGGAGGTGCCGCCACCGGCACCGCGGAGCTT
TTTTCTGTAGAGCATTGTGCCTATTTCCCCGAGTTTTTGCTGCCGAAGCTGTGACTGCCGAT
TCGGAAGTCCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGTCCCAGAGCCCTATTACCCGGA
ATTTGGATGGGACCGCCTCCGGGAGCTGTTTGGCAAAGATGAACAGCAGAGAATTTCAAAGG
ACCTTGCTGATATNTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGGTGTATGGGGGA
ATACCAGCTTTTATTCATGNTAAACAACAATACATTGAGCAGAGCCAGGCAGAAATTTATNA
TAACC

FIGURE 14

GAGCCGCCGCCGCGCGCGCCGCGCACTGCAGCCCCAGGCCCCGGCCCCCCACCCACGTCT
GCGTTGCTGCCCCGCCTGGGCCAGGCCCCAAAGGCAAGGACAAAGCAGCTGTCAGGGAACCT
CCGCCGGAGTCGAATTTACGTGCAGCTGCCGGCAACCACAGGTTCCAAGATGGTTTGCGGGG
GCTTCGCGTGTTCCAAGAACTGCCTGTGCGCCCTCAACCTGCTTTACACCTTGGTTAGTCTG
CTGCTAATTGGAATTGCTGCGTGGGGCATTGGCTTCGGGCTGATTTCCAGTCTCCGAGTGGT
CGGCGTGGTCATTGCAGTGGGCATCTTCTTGTTCCTGATTGCTTTAGTGGGTCTGATTGGAG
CTGTAAAACATCATCAGGTGTTGCTATTTTTTTATATGATTATTCTGTTACTTGTATTTATT
GTTCAGTTTTCTGTATCTTGCGCTTGTTTAGCCCTGAACCAGGAGCAACAGGGTCAGCTTCT
GGAGGTTGGTTGGAACAATACGGCAAGTGCTCGAAATGACATCCAGAGAAATCTAAACTGCT
GTGGGTTCCGAAGTGTTAACCCAAATGACACCTGTCTGGCTAGCTGTGTTAAAAGTGACCAC
TCGTGCTCGCCATGTGCTCCAATCATAGGAGAATATGCTGGAGAGGTTTTGAGATTTGTTGG
TGGCATTGGCCTGTTCTTCAGTTTTACAGAGATCCTGGGTGTTTGGCTGACCTACAGATACA
GGAACCAGAAAGACCCCCGCGCGAATCCTAGTGCATTCCTTTGATGAGAAAACAAGGAAGAT
TTCCTTTCGTATTATGATCTTGTTCACTTTCTGTAATTTTCTGTTAAGCTCCATTTGCCAGT
TTAAGGAAGGAAACACTATCTGGAAAAGTACCTTATTGATAGTGGAATTATATATTTTTACT
CTATGTTTCTCTACATGTTTTTTTCTTTCCGTTGCTGAAAAATATTTGAAACTTGTGGTCTC
TGAAGCTCGGTGGCACCTGGAATTTACTGTATTCATTGTCGGGCACTGTCCACTGTGGCCTT
TCTTAGCATTTTTACCTGCAGAAAAACTTTGTATGGTACCACTGTGTTGGTTATATGGTGAA
TCTGAACGTACATCTCACTGGTATAATTATATGTAGCACTGTGCTGTGTAGATAGTTCCTAC
TGGAAAAAGAGTGGAAATTTATTAAAATCAGAAAGTATGAGATCCTGTTATGTTAAGGGAAA
TCCAAATTCCCAATTTTTTTTGGTCTTTTTAGGAAAGATTGTTGTGGTAAAAAGTGTTAGTA
TAAAAATGATAATTTACTTGTAGTCTTTTATGATTACACCAATGTATTCTAGAAATAGTTAT
GTCTTAGGAAATTGTGGTTTAATTTTTGACTTTTACAGGTAAGTGCAAAGGAGAAGTGGTTT
CATGAAATGTTCTAATGTATAATAACATTTACCTTCAGCCTCCATCAGAATGGAACGAGTTT
TGAGTAATCAGGAAGTATATCTATATGATCTTGATATTGTTTTATAATAATTTGAAGTCTAA
AAGACTGCATTTTTAAACAAGTTAGTATTAATGCGTTGGCCCACGTAGCAAAAAGATATTTG
ATTATCTTAAAAATTGTTAAATACCGTTTTCATGAAATTTCTCAGTATTGTAACAGCAACTT
GTCAAACCTAAGCATATTTGAATATGATCTCCCATAATTTGAAATTGAAATCGTATTGTGTG
GCTCTGTATATTCTGTTAAAAAATTAAAGGACAGAAACCTTTCTTTGTGTATGCATGTTTGA
ATTAAAAGAAAGTAATGGAAG

FIGURE 15

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA39979
><subunit 1 of 1, 204 aa, 1 stop
><MW: 22147, pI: 8.37, NX(S/T): 3
MVCGGFACSKNCLCALNLLYTLVSLLLIGIAAWGIGFGLISSLRVVGVVIAVGIFLFLIALV
GLIGAVKHHQVLLFFYMIILLLVFIVQFSVSCACLALNQEQQGQLLEVGWNNTASARNDIQR
NLNCCGFRSVNPNDTCLASCVKSDHSCSPCAPIIGEYAGEVLRFVGGIGLFFSFTEILGVWL
TYRYRNQKDPRANPSAFL

Signal Peptide:
amino acids 1-34

Transmembrane domains:
amino acids 47-63, 72-95 and 162-182

FIGURE 16

TGATTGGAGCTGTAAAAAANTCTTCAGGTGTTGTNATTTTTTTATATGATTATTCTGTAANT
TGTATTTATTGTTCAGTTTTNTGTATCTTGCGCTTGTTTAGCCNTGAACCAGGAGCAACAGG
GTCAGNTTNTGGAGGTTGGTTGGAACAATACGGCAAGTGCTCGAAATGACATCCAGAGAAAT
NTAAACTGCTGTGGGTTCCGAAGTGTTAACCCAAATGACACCTGTNTGGCTAGCTGTGTTAA
AAGTGACCACTNGTGCTCGCCATGTGCTCCAATCATAGGAGAATATGCTGGAGAGGTTTTGA
GATTTGTTGGTGGCATTGGCCTGTTNTTCAGTTTTACAGAGATCCTGGGTGTTTGGCTGACC
TACAGATACAGGAACCAG

FIGURE 17

AATCCCAAATTCCCCAATTTTTTTGGNCTTTTTAGGGAAAGATGTGTTGTGGTAAAAAGTGT
TAGTATAAAAATGATAATTTACTTGTAGTCTTTTATGATTACACCAATGTATTCTAGAATAG
TTATGTCTTAGGAAATTGTGGTTTAATTTTTGACTTTTACAGGTAAGTGCAAAGGAGAAGTG
GTTTCATGAAATGTTCTAATGTATAATAACATTTACCTTCAGCCTCCCATCAGAATGGAACG
AGTTTTGAGTAATCCAGGAAGTATATCTATATGATCTTGATATTGTTTTATATAATTTGAAG
TCTAAAAGACTGCATTTTTAAACAAGTTAGTATTAATGCGTTGGCCCACGTAGCAAAAAGAT
ATTTGATTATCTTAAAAATTGTTAAATACCGTTTTCATGAAAGTTCTCAGTATTGTAACAGC
AACTTGTCAAACCTAAGCATATTTGAATATGATCTCCCATAATTTGAAATTGAAATCGTATT
GTGTGGAGGAAATGGCAATCTTATGTGTGCTGAAGGACACAGTAAGAGCACCAAGTTGTGCC
CCACTTGC

FIGURE 18

ATGATTATTCTGTTACTTGTATTTATTGTTCAGTTTTATGGTATCTTGCGCTTGTTTAGCCC
CTGAAACCAGGAGCAACAGGGNNCAGCTTCCTGGAGGTTGGTTGGCAACAATCACGGCCAAG
TGACTCCGCAAATGACATCCCAGAGAAATCCTAAACTGCTGTGGGTTCCGAAGTGTTAACCC
AAATGACACCTGTCTGGCTNGCTGTGTTAAAAGTGACCACTCGTGCTCGCCATGTGCTCCAA
TCATAGGAGAATATGC

FIGURE 19

CAGTCACCATGAAGCTGGGCTGTGTCCTCATGGCCTGGGCCCTCTACCTTTCCCTTGGTGTG
CTCTGGGTGGCCCAGATGCTACTGGCTGCCAGTTTTGAGACGCTGCAGTGTGAGGGACCTGT
CTGCACTGAGGAGAGCAGCTGCCACACGGAGGATGACTTGACTGATGCAAGGGAAGCTGGCT
TCCAGGTCAAGGCCTACACTTTCAGTGAACCCTTCCACCTGATTGTGTCCTATGACTGGCTG
ATCCTCCAAGGTCCAGCCAAGCCAGTTTTTGAAGGGGACCTGCTGGTTCTGCGCTGCCAGGC
CTGGCAAGACTGGCCACTGACTCAGGTGACCTTCTACCGAGATGGCTCAGCTCTGGGTCCCC
CCGGGCCTAACAGGGAATTCTCCATCACCGTGGTACAAAAGGCAGACAGCGGGCACTACCAC
TGCAGTGGCATCTTCCAGAGCCCTGGTCCTGGGATCCCAGAAACAGCATCTGTTGTGGCTAT
CACAGTCCAAGAACTGTTTCCAGCGCCAATTCTCAGAGCTGTACCCTCAGCTGAACCCCAAG
CAGGAAGCCCCATGACCCTGAGTTGTCAGACAAAGTTGCCCCTGCAGAGGTCAGCTGCCCGC
CTCCTCTTCTCCTTCTACAAGGATGGAAGGATAGTGCAAAGCAGGGGGCTCTCCTCAGAATT
CCAGATCCCCACAGCTTCAGAAGATCACTCCGGGTCATACTGGTGTGAGGCAGCCACTGAGG
ACAACCAAGTTTGGAAACAGAGCCCCCAGCTAGAGATCAGAGTGCAGGGTGCTTCCAGCTCT
GCTGCACCTCCCACATTGAATCCAGCTCCTCAGAAATCAGCTGCTCCAGGAACTGCTCCTGA
GGAGGCCCCTGGGCCTCTGCCTCCGCCGCCAACCCCATCTTCTGAGGATCCAGGCTTTTCTT
CTCCTCTGGGGATGCCAGATCCTCATCTGTATCACCAGATGGGCCTTCTTCTCAAACACATG
CAGGATGTGAGAGTCCTCCTCGGTCACCTGCTCATGGAGTTGAGGGAATTATCTGGCCACCA
GAAGCCTGGGACCACAAAGGCTACTGCTGAATAGAAGTAAACAGTTCATCCATGATCTCACT
TAACCACCCCAATAAATCTGATTCTTTATTTTCTCTTCCTGTCCTGCACATATGCATAAGTA
CTTTTACAAGTTGTCCCAGTGTTTTGTTAGAATAATGTAGTTAGGTGAGTGTAAATAAATTT
ATATAAAGTGAGAATTAGAGTTTAGCTATAATTGTGTATTCTCTCTTAACACAACAGAATTC
TGCTGTCTAGATCAGGAATTTCTATCTGTTATATCGACCAGAATGTTGTGATTTAAAGAGAA
CTAATGGAAGTGGATTGAATACAGCAGTCTCAACTGGGGGCAATTTTGCCCCCCAGAGGACA
TTGGGCAATGTTTGGAGACATTTTGGTCATTATACTTGGGGGGTTGGGGGATGGTGGGATGT
GTGTCTACTGGCATCCAGTAAATAGAAGCCAGGGGTGCCGCTAAACATCCTATAATGCACAG
GGCAGTACCCCACAACGAAAAATAATCTGGCCCAAAATGTCAGTTGTACTGAGTTTGAGAAA
CCCCAGCCTAATGAAACCCTAGGTGTTGGGCTCTGGAATGGGACTTTGTCCCTTCTAATTAT
TATCTCTTTCCAGCCTCATTCAGCTATTCTTACTGACATACCAGTCTTTAGCTGGTGCTATG
GTCTGTTCTTTAGTTCTAGTTTGTATCCCCTCAAAAGCCATTATGTTGAAATCCTAATCCCC
AAGGTGATGGCATTAAGAAGTGGGCCTTTGGGAAGTGATTAGATCAGGAGTGCAGAGCCCTC
ATGATTAGGATTAGTGCCCTTATTTAAAAAGGCCCCAGAGAGCTAACTCACCCTTCCACCAT
ATGAGGACGTGGCAAGAAGATGACATGTATGAGAACCAAAAAACAGCTGTCGCCAAACACCG
ACTCTGTCGTTGCCTTGATCTTGAACTTCCAGCCTCCAGAACTATGAGAAATAAAATTCTGG
TTGTTTGTAGCCTAA

FIGURE 20

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA40594

><subunit 1 of 1, 359 aa, 1 stop

><MW: 38899, pI: 5.21, NX(S/T): 0

MKLGCVLMAWALYLSLGVLWVAQMLLAASFETLQCEGPVCTEESSCHTEDDLTDAREAGFQV
KAYTFSEPFHLIVSYDWLILQGPAKPVFEGDLLVLRCQAWQDWPLTQVTFYRDGSALGPPGP
NREFSITVVQKADSGHYHCSGIFQSPGPGIPETASVVAITVQELFPAPILRAVPSAEPQAGS
PMTLSCQTKLPLQRSAARLLFSFYKDGRIVQSRGLSSEFQIPTASEDHSGSYWCEAATEDNQ
VWKQSPQLEIRVQGASSSAAPPTLNPAPQKSAAPGTAPEEAPGPLPPPPTPSSEDPGFSSPL
GMPDPHLYHQMGLLLKHMQDVRVLLGHLLMELRELSGHQKPGTTKATAE

Leucine zipper pattern sequence:

amino acids 12-33

Protein kinase C phosphorylation site:

amino acids 353-355

FIGURE 21

CCCACGCGTCCGCCCACGCGTCCGCCCACGGGTCCGCCCACGCGTCCGGGCCACCAGAAGTT
TGAGCCTCTTTGGTAGCAGGAGGCTGGAAGAAAGGACAGAAGTAGCTCTGGCTGTGATGGGG
ATCTTACTGGGCCTGCTACTCCTGGGGCACCTAACAGTGGACACTTATGGCCGTCCCATCCT
GGAAGTGCCAGAGAGTGTAACAGGACCTTGGAAAGGGGATGTGAATCTTCCCTGCACCTATG
ACCCCCTGCAAGGCTACACCCAAGTCTTGGTGAAGTGGCTGGTACAACGTGGCTCAGACCCT
GTCACCATCTTTCTACGTGACTCTTCTGGAGACCATATCCAGCAGGCAAAGTACCAGGGCCG
CCTGCATGTGAGCCACAAGGTTCCAGGAGATGTATCCCTCCAATTGAGCACCCTGGAGATGG
ATGACCGGAGCCACTACACGTGTGAAGTCACCTGGCAGACTCCTGATGGCAACCAAGTCGTG
AGAGATAAGATTACTGAGCTCCGTGTCCAGAAACTCTCTGTCTCCAAGCCCACAGTGACAAC
TGGCAGCGGTTATGGCTTCACGGTGCCCCAGGGAATGAGGATTAGCCTTCAATGCCAGGCTC
GGGGTTCTCCTCCCATCAGTTATATTTGGTATAAGCAACAGACTAATAACCAGGAACCCATC
AAAGTAGCAACCCTAAGTACCTTACTCTTCAAGCCTGCGGTGATAGCCGACTCAGGCTCCTA
TTTCTGCACTGCCAAGGGCCAGGTTGGCTCTGAGCAGCACAGCGACATTGTGAAGTTTGTGG
TCAAAGACTCCTCAAAGCTACTCAAGACCAAGACTGAGGCACCTACAACCATGACATACCCC
TTGAAAGCAACATCTACAGTGAAGCAGTCCTGGGACTGGACCACTGACATGGATGGCTACCT
TGGAGAGACCAGTGCTGGGCCAGGAAAGAGCCTGCCTGTCTTTGCCATCATCCTCATCATCT
CCTTGTGCTGTATGGTGGTTTTTACCATGGCCTATATCATGCTCTGTCGGAAGACATCCCAA
CAAGAGCATGTCTACGAAGCAGCCAGGTAAGAAAGTCTCTCCTCTTCCATTTTTGACCCCGT
CCCTGCCCTCAATTTTGATTACTGGCAGGAAATGTGGAGGAAGGGGGGTGTGGCACAGACCC
AATCCTAAGGCCGGAGGCCTTCAGGGTCAGGACATAGCTGCCTTCCCTCTCTCAGGCACCTT
CTGAGGTTGTTTTGGCCCTCTGAACACAAAGGATAATTTAGATCCATCTGCCTTCTGCTTCC
AGAATCCCTGGGTGGTAGGATCCTGATAATTAATTGGCAAGAATTGAGGCAGAAGGGTGGGA
AACCAGGACCACAGCCCCAAGTCCCTTCTTATGGGTGGTGGGCTCTTGGGCCATAGGGCACA
TGCCAGAGAGGCCAACGACTCTGGAGAAACCATGAGGGTGGCCATCTTCGCAAGTGGCTGCT
CCAGTGATGAGCCAACTTCCCAGAATCTGGGCAACAACTACTCTGATGAGCCCTGCATAGGA
CAGGAGTACCAGATCATCGCCCAGATCAATGGCAACTACGCCCGCCTGCTGGACACAGTTCC
TCTGGATTATGAGTTTCTGGCCACTGAGGGCAAAAGTGTCTGTTAAAAATGCCCCATTAGGC
CAGGATCTGCTGACATAATTGCCTAGTCAGTCCTTGCCTTCTGCATGGCCTTCTTCCCTGCT
ACCTCTCTTCCTGGATAGCCCAAAGTGTCCGCCTACCAACACTGGAGCCGCTGGGAGTCACT
GGCTTTGCCCTGGAATTTGCCAGATGCATCTCAAGTAAGCCAGCTGCTGGATTTGGCTCTGG
GCCCTTCTAGTATCTCTGCCGGGGGCTTCTGGTACTCCTCTCTAAATACCAGAGGGAAGATG
CCCATAGCACTAGGACTTGGTCATCATGCCTACAGACACTATTCAACTTTGGCATCTTGCCA
CCAGAAGACCCGAGGGAGGCTCAGCTCTGCCAGCTCAGAGGACCAGCTATATCCAGGATCAT
TTCTCTTTCTTCAGGGCCAGACAGCTTTTAATTGAAATTGTTATTTCACAGGCCAGGGTTCA
GTTCTGCTCCTCCACTATAAGTCTAATGTTCTGACTCTCTCCTGGTGCTCAATAAATATCTA
ATCATAACAGC

FIGURE 22

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45416
><subunit 1 of 1, 321 aa, 1 stop
><MW: 35544, pI: 8.51, NX(S/T): 0
MGILLGLLLLGHLTVDTYGRPILEVPESVTGPWKGDVNLPCTYDPLQGYTQVLVKWLVQRGS
DPVTIFLRDSSGDHIQQAKYQGRLHVSHKVPGDVSLQLSTLEMDDRSHYTCEVTWQTPDGNQ
VVRDKITELRVQKLSVSKPTVTTGSGYGFTVPQGMRISLQCQARGSPPISYIWYKQQTNNQE
PIKVATLSTLLFKPAVIADSGSYFCTAKGQVGSEQHSDIVKFVVKDSSKLLKTKTEAPTTMT
YPLKATSTVKQSWDWTTDMDGYLGETSAGPGKSLPVFAIILIISLCCMVVFTMAYIMLCRKT
SQQEHVYEAAR

Glycosaminoglycan attachment site:
amino acids 149-152

Transmembrane domain:
amino acids 276-306

FIGURE 23

GCGCCGGGAGCCCATCTGCCCCCAGGGGCACGGGGCGCGGGGCCGGCTCCCGCCCGGCACAT
GGCTGCAGCCACCTCGCGCGCACCCCGAGGCGCCGCGCCCAGCTCGCCCGAGGTCCGTCGGA
GGCGCCCGGCCGCCCCCGGAGCCAAGCAGCAACTGAGCGGGGAAGCGCCCGCGTCCGGGGATC
GGGATGTCCCTCCTCCTTCTCCTCTTGCTAGTTTCCTACTATGTTGGAACCTTGGGGACTCA
CACTGAGATCAAGAGAGTGGCAGAGGAAAAGGTCACTTTGCCCTGCCACCATCAACTGGGGC
TTCCAGAAAAAGACACTCTGGATATTGAATGGCTGCTCACCGATAATGAAGGGAACCAAAAA
GTGGTGATCACTTACTCCAGTCGTCATGTCTACAATAACTTGACTGAGGAACAGAAGGGCCG
AGTGGCCTTTGCTTCCAATTTCCTGGCAGGAGATGCCTCCTTGCAGATTGAACCTCTGAAGC
CCAGTGATGAGGGCCGGTACACCTGTAAGGTTAAGAATTCAGGGCGCTACGTGTGGAGCCAT
GTCATCTTAAAAGTCTTAGTGAGACCATCCAAGCCCAAGTGTGAGTTGGAAGGAGAGCTGAC
AGAAGGAAGTGACCTGACTTTGCAGTGTGAGTCATCCTCTGGCACAGAGCCCATTGTGTATT
ACTGGCAGCGAATCCGAGAGAAAGAGGGAGAGGATGAACGTCTGCCTCCCAAATCTAGGATT
GACTACAACCACCCTGGACGAGTTCTGCTGCAGAATCTTACCATGTCCTACTCTGGACTGTA
CCAGTGCACAGCAGGCAACGAAGCTGGGAAGGAAAGCTGTGTGGTGCGAGTAACTGTACAGT
ATGTACAAAGCATCGGCATGGTTGCAGGAGCAGTGACAGGCATAGTGGCTGGAGCCCTGCTG
ATTTTCCTCTTGGTGTGGCTGCTAATCCGAAGGAAAGACAAAGAAAGATATGAGGAAGAAGA
GAGACCTAATGAAATTCGAGAAGATGCTGAAGCTCCAAAAGCCCGTCTTGTGAAACCCAGCT
CCTCTTCCTCAGGCTCTCGGAGCTCACGCTCTGGTTCTTCCTCCACTCGCTCCACAGCAAAT
AGTGCCTCACGCAGCCAGCGGACACTGTCAACTGACGCAGCACCCCAGCCAGGGCTGGCCAC
CCAGGCATACAGCCTAGTGGGGCCAGAGGTGAGAGGTTCTGAACCAAAGAAAGTCCACCATG
CTAATCTGACCAAAGCAGAAACCACACCCAGCATGATCCCCAGCCAGAGCAGAGCCTTCCAA
ACGGTCTGAATTACAATGGACTTGACTCCCACGCTTTCCTAGGAGTCAGGGTCTTTGGACTC
TTCTCGTCATTGGAGCTCAAGTCACCAGCCACACAACCAGATGAGAGGTCATCTAAGTAGCA
GTGAGCATTGCACGGAACAGATTCAGATGAGCATTTTCCTTATACAATACCAAACAAGCAAA
AGGATGTAAGCTGATTCATCTGTAAAAAGGCATCTTATTGTGCCTTTAGACCAGAGTAAGGG
AAAGCAGGAGTCCAAATCTATTTGTTGACCAGGACCTGTGGTGAGAAGGTTGGGGAAAGGTG
AGGTGAATATACCTAAAACTTTTAATGTGGGATATTTTGTATCAGTGCTTTGATTCACAATT
TTCAAGAGGAAATGGGATGCTGTTTGTAAATTTTCTATGCATTTCTGCAAACTTATTGGATT
ATTAGTTATTCAGACAGTCAAGCAGAACCCACAGCCTTATTACACCTGTCTACACCATGTAC
TGAGCTAACCACTTCTAAGAAACTCCAAAAAAGGAAACATGTGTCTTCTATTCTGACTTAAC
TTCATTTGTCATAAGGTTTGGATATTAATTTCAAGGGGAGTTGAAATAGTGGGAGATGGAGA
AGAGTGAATGAGTTTCTCCCACTCTATACTAATCTCACTATTTGTATTGAGCCCAAAATAAC
TATGAAAGGAGACAAAAATTTGTGACAAAGGATTGTGAAGAGCTTTCCATCTTCATGATGTT
ATGAGGATTGTTGACAAACATTAGAAATATATAATGGAGCAATTGTGGATTTCCCCTCAAAT
CAGATGCCTCTAAGGACTTTCCTGCTAGATATTTCTGGAAGGAGAAAATACAACATGTCATT
TATCAACGTCCTTAGAAAGAATTCTTCTAGAGAAAAAGGGATCTAGGAATGCTGAAAGATTA
CCCAACATACCATTATAGTCTCTTCTTTCTGAGAAAATGTGAAACCAGAATTGCAAGACTGG
GTGGACTAGAAAGGGAGATTAGATCAGTTTTCTCTTAATATGTCAAGGAAGGTAGCCGGGCA
TGGTGCCAGGCACCTGTAGGAAAATCCAGCAGGTGGAGGTTGCAGTGAGCCGAGATTATGCC
ATTGCACTCCAGCCTGGGTGACAGAGCGGGACTCCGTCTC

FIGURE 24

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45419
><subunit 1 of 1, 373 aa, 1 stop
><MW: 41281, pI: 8.33, NX(S/T): 3
MSLLLLLLLVSYYVGTLGTHTEIKRVAEEKVTLPCHHQLGLPEKDTLDIEWLLTDNEGNQKV
VITYSSRHVYNNLTEEQKGRVAFASNFLAGDASLQIEPLKPSDEGRYTCKVKNSGRYVWSHV
ILKVLVRPSKPKCELEGELTEGSDLTLQCESSSGTEPIVYYWQRIREKEGEDERLPPKSRID
YNHPGRVLLQNLTMSYSGLYQCTAGNEAGKESCVVRVTVQYVQSIGMVAGAVTGIVAGALLI
FLLVWLLIRRKDKERYEEEERPNEIREDAEAPKARLVKPSSSSSGSRSSRSGSSSTRSTANS
ASRSQRTLSTDAAPQPGLATQAYSLVGPEVRGSEPKKVHHANLTKAETTPSMIPSQSRAFQTV

Transmembrane domain:
amino acids 221-254

FIGURE 25

GTCGTTCCTTTGCTCTCTCGCGCCCAGTCCTCCTCCCTGGTTCTCCTCAGCCGCTGTCGGAG
GAGAGCACCCGGAGACGCGGGCTGCAGTCGCGGCGGCTTCTCCCCGCCTGGGCGGCCTCGCC
GCTGGGCAGGTGCTGAGCGCCCCTAGAGCCTCCCTTGCCGCCTCCCTCCTCTGCCCGGCCGC
AGCAGTGCACATGGGGTGTTGGAGGTAGATGGGCTCCCGGCCCGGGAGGCGGCGGTGGATGC
GGCGCTGGGCAGAAGCAGCCGCCGATTCCAGCTGCCCCGCGCGCCCCGGGCGCCCCTGCGAG
TCCCCGGTTCAGCC**ATG**GGGACCTCTCCGAGCAGCAGCACCGCCCTCGCCTCCTGCAGCCGC
ATCGCCCGCCGAGCCACAGCCACGATGATCGCGGGCTCCCTTCTCCTGCTTGGATTCCTTAG
CACCACCACAGCTCAGCCAGAACAGAAGGCCTCGAATCTCATTGGCACATACCGCCATGTTG
ACCGTGCCACCGGCCAGGTGCTAACCTGTGACAAGTGTCCAGCAGGAACCTATGTCTCTGAG
CATTGTACCAACACAAGCCTGCGCGTCTGCAGCAGTTGCCCTGTGGGGACCTTTACCAGGCA
TGAGAATGGCATAGAGAAATGCCATGACTGTAGTCAGCCATGCCCATGGCCAATGATTGAGA
AATTACCTTGTGCTGCCTTGACTGACCGAGAATGCACTTGCCCACCTGGCATGTTCCAGTCT
AACGCTACCTGTGCCCCCCATACGGTGTGTCCTGTGGGTTGGGGTGTGCGGAAGAAAGGGAC
AGAGACTGAGGATGTGCGGTGTAAGCAGTGTGCTCGGGGTACCTTCTCAGATGTGCCTTCTA
GTGTGATGAAATGCAAAGCATACACAGACTGTCTGAGTCAGAACCTGGTGGTGATCAAGCCG
GGGACCAAGGAGACAGACAACGTCTGTGGCACACTCCCGTCCTTCTCCAGCTCCACCTCACC
TTCCCCTGGCACAGCCATCTTTCCACGCCCTGAGCACATGGAAACCCATGAAGTCCCTTCCT
CCACTTATGTTCCCAAAGGCATGAACTCAACAGAATCCAACTCTTCTGCCTCTGTTAGACCA
AAGGTACTGAGTAGCATCCAGGAAGGGACAGTCCCTGACAACACAAGCTCAGCAAGGGGGAA
GGAAGACGTGAACAAGACCCTCCCAAACCTTCAGGTAGTCAACCACCAGCAAGGCCCCCACC
ACAGACACATCCTGAAGCTGCTGCCGTCCATGGAGGCCACTGGGGGCGAGAAGTCCAGCACG
CCCATCAAGGGCCCCAAGAGGGGACATCCTAGACAGAACCTACACAAGCATTTTGACATCAA
TGAGCATTTGCCCTGGATGATTGTGCTTTTCCTGCTGCTGGTGCTTGTGGTGATTGTGGTGT
GCAGTATCCGGAAAAGCTCGAGGACTCTGAAAAAGGGGCCCCGGCAGGATCCCAGTGCCATT
GTGGAAAAGGCAGGGCTGAAGAAATCCATGACTCCAACCCAGAACCGGGAGAAATGGATCTA
CTACTGCAATGGCCATGGTATCGATATCCTGAAGCTTGTAGCAGCCCAAGTGGGAAGCCAGT
GGAAAGATATCTATCAGTTTCTTTGCAATGCCAGTGAGAGGGAGGTTGCTGCTTTCTCCAAT
GGGTACACAGCCGACCACGAGCGGGCCTACGCAGCTCTGCAGCACTGGACCATCCGGGGCCC
CGAGGCCAGCCTCGCCCAGCTAATTAGCGCCCTGCGCCAGCACCGGAGAAACGATGTTGTGG
AGAAGATTCGTGGGCTGATGGAAGACACCACCCAGCTGGAAACTGACAAACTAGCTCTCCCG
ATGAGCCCCAGCCCGCTTAGCCCGAGCCCCATCCCCAGCCCCAACGCGAAACTTGAGAATTC
CGCTCTCCTGACGGTGGAGCCTTCCCCACAGGACAAGAACAAGGGCTTCTTCGTGGATGAGT
CGGAGCCCCTTCTCCGCTGTGACTCTACATCCAGCGGCTCCTCCGCGCTGAGCAGGAACGGT
TCCTTTATTACCAAAGAAAAGAAGGACACAGTGTTGCGGCAGGTACGCCTGGACCCCTGTGA
CTTGCAGCCTATCTTTGATGACATGCTCCACTTTCTAAATCCTGAGGAGCTGCGGGTGATTG
AAGAGATTCCCCAGGCTGAGGACAAACTAGACCGGCTATTCGAAATTATTGGAGTCAAGAGC
CAGGAAGCCAGCCAGACCCTCCTGGACTCTGTTTATAGCCATCTTCCTGACCTGCTG**TAG**AA
CATAGGGATACTGCATTCTGGAAATTACTCAATTTAGTGGCAGGGTGGTTTTTTAATTTTCT
TCTGTTTCTGATTTTTGTTGTTTGGGGTGTGTGTGTGTGTTTGTGTGTGTGTGTGTGTGT
GTGTGTGTGTGTGTTTAACAGAGAATATGGCCAGTGCTTGAGTTCTTTCTCCTTCTCTCTCT
CTCTTTTTTTTTTAAATAACTCTTCTGGGAAGTTGGTTTATAAGCCTTTGCCAGGTGTAACT
GTTGTGAAATACCCACCACTAAAGTTTTTTAAGTTCCATATTTTCTCCATTTTGCCTTCTTA
TGTATTTTCAAGATTATTCTGTGCACTTTAAATTTACTTAACTTACCATAAATGCAGTGTGA
CTTTTCCCACACACTGGATTGTGAGGCTCTTAACTTCTTAAAAGTATAATGGCATCTTGTGA
ATCCTATAAGCAGTCTTTATGTCTCTTAACATTCACACCTACTTTTTAAAAACAAATATTAT
TACTATTTTTATTATTGTTTGTCCTTTATAAATTTTCTTAAAGATTAAGAAAATTTAAGACC
CCATTGAGTTACTGTAATGCAATTCAACTTTGAGTTATCTTTTAAATATGTCTTGTATAGTT
CATATTCATGGCTGAAACTTGACCACACTATTGCTGATTGTATGGTTTTCACCTGGACACCG
TGTAGAATGCTTGATTACTTGTACTCTTCTTATGCTAATATGCTCTGGGCTGGAGAAATGAA
ATCCTCAAGCCATCAGGATTTGCTATTTAAGTGGCTTGACAACTGGGCCACCAAAGAACTTG
AACTTCACCTTTTAGGATTTGAGCTGTTCTGGAACACATTGCTGCACTTTGGAAAGTCAAAA
TCAAGTGCCAGTGGCGCCCTTTCCATAGAGAATTTGCCCAGCTTTGCTTTAAAAGATGTCTT
GTTTTTTATATACACATAATCAATAGGTCCAATCTGCTCTCAAGGCCTTGGTCCTGGTGGGA
TTCCTTCACCAATTACTTTAATTAAAAATGGCTGCAACTGTAAGAACCCTTGTCTGATATAT
TTGCAACTATGCTCCCATTTACAAATGTACCTTCTAATGCTCAGTTGCCAGGTTCCAATGCA
AAGGTGGCGTGGACTCCCTTTGTGTGGGTGGGGTTTGTGGGTAGTGGTGAAGGACCGATATC
AGAAAAATGCCTTCAAGTGTACTAATTTATTAATAAACATTAGGTGTTTGTTAAAAAAAAAA

FIGURE 26

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52594
><subunit 1 of 1, 655 aa, 1 stop
><MW: 71845, pI: 8.22, NX(S/T): 8
MGTSPSSSTALASCSRIARRATATMIAGSLLLLGFLSTTTAQPEQKASNLIGTYRHVDRATG
QVLTCDKCPAGTYVSEHCTNTSLRVCSSCPVGTFTRHENGIEKCHDCSQPCPWPMIEKLPCA
ALTDRECTCPPGMFQSNATCAPHTVCPVGWGVRKKGTETEDVRCKQCARGTFSDVPSSVMKC
KAYTDCLSQNLVVIKPGTKETDNVCGTLPSFSSSTSPSPGTAIFPRPEHMETHEVPSSTYVP
KGMNSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHIL
KLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCSIRK
SSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQWKDIY
QFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRNDVVEKIRG
LMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLTVEPSPQDKNKGFFVDESEPLL
RCDSTSSGSSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHFLNPEELRVIEEIPQ
AEDKLDRLFEIIGVKSQEASQTLLDSVYSHLPDLL

FIGURE 27

ATGGGAAGCCAGTAACACTGTGGCCTACTATCTCTTCCGTGGTGCCATCTACATTTTTGGGA
CTCGGGAATTATGAGGTAGAGGTGGAGGCGGAGCCGGATGTCAGAGGTCCTGAAATAGTCAC
CATGGGGGAAAATGATCCGCCTGCTGTTGAAGCCCCCTTCTCATTCCGATCGCTTTTTGGCC
TTGATGATTTGAAAATAAGTCCTGTTGCACCAGATGCAGATGCTGTTGCTGCACAGATCCTG
TCACTGCTGCCATTGAAGTTTTTTCCAATCATCGTCATTGGGATCATTGCATTGATATTAGC
ACTGGCCATTGGTCTGGGCATCCACTTCGACTGCTCAGGGAAGTACAGATGTCGCTCATCCT
TTAAGTGTATCGAGCTGATAGCTCGATGTGACGGAGTCTCGGATTGCAAAGACGGGGAGGAC
GAGTACCGCTGTGTCCGGGTGGGTGGTCAGAATGCCGTGCTCCAGGTGTTCACAGCTGCTTC
GTGGAAGACCATGTGCTCCGATGACTGGAAGGGTCACTACGCAAATGTTGCCTGTGCCCAAC
TGGGTTTCCCAAGCTATGTGAGTTCAGATAACCTCAGAGTGAGCTCGCTGGAGGGGCAGTTC
CGGGAGGAGTTTGTGTCCATCGATCACCTCTTGCCAGATGACAAGGTGACTGCATTACACCA
CTCAGTATATGTGAGGGAGGGATGTGCCTCTGGCCACGTGGTTACCTTGCAGTGCACAGCCT
GTGGTCATAGAAGGGGCTACAGCTCACGCATCGTGGGTGGAAACATGTCCTTGCTCTCGCAG
TGGCCCTGGCAGGCCAGCCTTCAGTTCCAGGGCTACCACCTGTGCGGGGGCTCTGTCATCAC
GCCCCTGTGGATCATCACTGCTGCACACTGTGTTTATGACTTGTACCTCCCCAAGTCATGGA
CCATCCAGGTGGGTCTAGTTTCCCTGTTGGACAATCCAGCCCCATCCCACTTGGTGGAGAAG
ATTGTCTACCACAGCAAGTACAAGCCAAAGAGGCTGGGCAATGACATCGCCCTTATGAAGCT
GGCCGGGCCACTCACGTTCAATGAAATGATCCAGCCTGTGTGCCTGCCCAACTCTGAAGAGA
ACTTCCCCGATGGAAAAGTGTGCTGGACGTCAGGATGGGGGGCCACAGAGGATGGAGGTGAC
GCCTCCCCTGTCCTGAACCACGCGGCCGTCCCTTTGATTTCCAACAAGATCTGCAACCACAG
GGACGTGTACGGTGGCATCATCTCCCCCTCCATGCTCTGCGCGGGCTACCTGACGGGTGGCG
TGGACAGCTGCCAGGGGGACAGCGGGGGGCCCCTGGTGTGTCAAGAGAGGAGGCTGTGGAAG
TTAGTGGGAGCGACCAGCTTTGGCATCGGCTGCGCAGAGGTGAACAAGCCTGGGGTGTACAC
CCGTGTCACCTCCTTCCTGGACTGGATCCACGAGCAGATGGAGAGAGACCTAAAAACCTGAA
GAGGAAGGGGACAAGTAGCCACCTGAGTTCCTGAGGTGATGAAGACAGCCCGATCCTCCCCT
GGACTCCCGTGTAGGAACCTGCACACGAGCAGACACCCTTGGAGCTCTGAGTTCCGGCACCA
GTAGCAGGCCCGAAAGAGGCACCCTTCCATCTGATTCCAGCACAACCTTCAAGCTGCTTTTT
GTTTTTTGTTTTTTTGAGGTGGAGTCTCGCTCTGTTGCCCAGGCTGGAGTGCAGTGGCGAAA
TCCCTGCTCACTGCAGCCTCCGCTTCCCTGGTTCAAGCGATTCTCTTGCCTCAGCTTCCCCA
GTAGCTGGGACCACAGGTGCCCGCCACCACACCCAACTAATTTTTGTATTTTTAGTAGAGAC
AGGGTTTCACCATGTTGGCCAGGCTGCTCTCAAACCCCTGACCTCAAATGATGTGCCTGCTT
CAGCCTCCCACAGTGCTGGGATTACAGGCATGGGCCACCACGCCTAGCCTCACGCTCCTTTC
TGATCTTCACTAAGAACAAAAGAAGCAGCAACTTGCAAGGGCGGCCTTTCCCACTGGTCCAT
CTGGTTTTCTCTCCAGGGTCTTGCAAAATTCCTGACGAGATAAGCAGTTATGTGACCTCACG
TGCAAAGCCACCAACAGCCACTCAGAAAAGACGCACCAGCCCAGAAGTGCAGAACTGCAGTC
ACTGCACGTTTTCATCTCTAGGGACCAGAACCAAACCCACCCTTTCTACTTCCAAGACTTAT
TTTCACATGTGGGGAGGTTAATCTAGGAATGACTCGTTTAAGGCCTATTTTCATGATTTCTT
TGTAGCATTTGGTGCTTGACGTATTATTGTCCTTTGATTCCAAATAATATGTTTCCTTCCCT
CATTGTCTGGCGTGTCTGCGTGGACTGGTGACGTGAATCAAAATCATCCACTGAAA

FIGURE 28

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45234
><subunit 1 of 1, 453 aa, 1 stop
><MW: 49334, pI: 6.32, NX(S/T): 1
MGENDPPAVEAPFSFRSLFGLDDLKISPVAPDADAVAAQILSLLPLKFFPIIVIGIIALILA
LAIGLGIHFDCSGKYRCRSSFKCIELIARCDGVSDCKDGEDEYRCVRVGGQNAVLQVFTAAS
WKTMCSDDWKGHYANVACAQLGFPSYVSSDNLRVSSLEGQFREEFVSIDHLLPDDKVTALHH
SVYVREGCASGHVVTLQCTACGHRRGYSSRIVGGNMSLLSQWPWQASLQFQGYHLCGGSVIT
PLWIITAAHCVYDLYLPKSWTIQVGLVSLLDNPAPSHLVEKIVYHSKYKPKRLGNDIALMKL
AGPLTFNEMIQPVCLPNSEENFPDGKVCWTSGWGATEDGGDASPVLNHAAVPLISNKICNHR
DVYGGIISPSMLCAGYLTGGVDSCQGDSGGPLVCQERRLWKLVGATSFGIGCAEVNKPGVYT
RVTSFLDWIHEQMERDLKT

FIGURE 29

CCCACGCGTCCGTCCTAGTCCCCGGGCCAACTCGGACAGTTTGCTCATTTATTGCAACGGTC
AAGGCTGGCTTGTGCCAGAACGGCGCGCGCGCGCGCACGCACGCACACACACGGGGGGAAAC
TTTTTTAAAAATGAAAGGCTAGAAGAGCTCAGCGGCGGCGCGGGCGCTGCGCGAGGGCTCCG
GAGCTGACTCGCCGAGGCAGGAAATCCCTCCGGTCGCGACGCCCGGCCCCGGCTCGGCGCCC
GCGTGGGATGGTGCAGCGCTCGCCGCCGGGCCCGAGAGCTGCTGCACTGAAGGCCGGCGACG
ATGGCAGCGCGCCCGCTGCCCGTGTCCCCCGCCCGCGCCCTCCTGCTCGCCCTGGCCGGTGC
TCTGCTCGCGCCCTGCGAGGCCCGAGGGGTGAGCTTATGGAACCAAGGAAGAGCTGATGAAG
TTGTCAGTGCCTCTGTTCGGAGTGGGGACCTCTGGATCCCAGTGAAGAGCTTCGACTCCAAG
AATCATCCAGAAGTGCTGAATATTCGACTACAACGGGAAAGCAAAGAACTGATCATAAATCT
GGAAAGAAATGAAGGTCTCATTGCCAGCAGTTTCACGGAAACCCACTATCTGCAAGACGGTA
CTGATGTCTCCCTCGCTCGAAATTACACGGGTCACTGTTACTACCATGGACATGTACGGGGA
TATTCTGATTCAGCAGTCAGTCTCAGCACGTGTTCTGGTCTCAGGGGACTTATTGTGTTTGA
AAATGAAAGCTATGTCTTAGAACCAATGAAAAGTGCAACCAACAGATACAAACTCTTCCCAG
CGAAGAAGCTGAAAAGCGTCCGGGGATCATGTGGATCACATCACAACACACCAAACCTCGCT
GCAAAGAATGTGTTTCCACCACCCTCTCAGACATGGGCAAGAAGGCATAAAAGAGAGACCCT
CAAGGCAACTAAGTATGTGGAGCTGGTGATCGTGGCAGACAACCGAGAGTTTCAGAGGCAAG
GAAAAGATCTGGAAAAAGTTAAGCAGCGATTAATAGAGATTGCTAATCACGTTGACAAGTTT
TACAGACCACTGAACATTCGGATCGTGTTGGTAGGCGTGGAAGTGTGGAATGACATGGACAA
ATGCTCTGTAAGTCAGGACCCATTCACCAGCCTCCATGAATTTCTGGACTGGAGGAAGATGA
AGCTTCTACCTCGCAAATCCCATGACAATGCGCAGCTTGTCAGTGGGGTTTATTTCCAAGGG
ACCACCATCGGCATGGCCCCAATCATGAGCATGTGCACGGCAGACCAGTCTGGGGGAATTGT
CATGGACCATTCAGACAATCCCCTTGGTGCAGCCGTGACCCTGGCACATGAGCTGGGCCACA
ATTTCGGGATGAATCATGACACACTGGACAGGGGCTGTAGCTGTCAAATGGCGGTTGAGAAA
GGAGGCTGCATCATGAACGCTTCCACCGGGTACCCATTTCCCATGGTGTTCAGCAGTTGCAG
CAGGAAGGACTTGGAGACCAGCCTGGAGAAAGGAATGGGGGTGTGCCTGTTTAACCTGCCGG
AAGTCAGGGAGTCTTTCGGGGGCCAGAAGTGTGGGAACAGATTTGTGGAAGAAGGAGAGGAG
TGTGACTGTGGGGAGCCAGAGGAATGTATGAATCGCTGCTGCAATGCCACCACCTGTACCCT
GAAGCCGGACGCTGTGTGCGCACATGGGCTGTGCTGTGAAGACTGCCAGCTGAAGCCTGCAG
GAACAGCGTGCAGGGACTCCAGCAACTCCTGTGACCTCCCAGAGTTCTGCACAGGGGCCAGC
CCTCACTGCCCAGCCAATGTGTACCTGCACGATGGGCACTCATGTCAGGATGTGGACGGCTA
CTGCTACAATGGCATCTGCCAGACTCACGAGCAGCAGTGTGTCACGCTCTGGGGACCAGGTG
CTAAACCTGCCCCTGGGATCTGCTTTGAGAGAGTCAATTCTGCAGGTGATCCTTATGGCAAC
TGTGGCAAAGTCTCGAAGAGTTCCTTTGCCAAATGCGAGATGAGAGATGCTAAATGTGGAAA
AATCCAGTGTCAAGGAGGTGCCAGCCGGCCAGTCATTGGTACCAATGCCGTTTCCATAGAAA
CAAACATCCCTCTGCAGCAAGGAGGCCGGATTCTGTGCCGGGGGACCCACGTGTACTTGGGC
GATGACATGCCGGACCCAGGGCTTGTGCTTGCAGGCACAAAGTGTGCAGATGGAAAAATCTG
CCTGAATCGTCAATGTCAAAATATTAGTGTCTTTGGGGTTCACGAGTGTGCAATGCAGTGCC
ACGGCAGAGGGGTGTGCAACAACAGGAAGAACTGCCACTGCGAGGCCCACTGGGCACCTCCC
TTCTGTGACAAGTTTGGCTTTGGAGGAAGCACAGACAGCGGCCCCATCCGGCAAGCAGAAGC
AAGGCAGGAAGCTGCAGAGTCCAACAGGGAGCGCGGCCAGGGCCAGGAGCCCGTGGGATCGC
AGGAGCATGCGTCTACTGCCTCACTGACACTCATC**TGA**GCCCTCCCATGACATGGAGACCGT
GACCAGTGCTGCTGCAGAGGAGGTCACGCGTCCCCAAGGCCTCCTGTGACTGGCAGCATTGA
CTCTGTGGCTTTGCCATCGTTTCCATGACAACAGACACAACACAGTTCTCGGGGCTCAGGAG
GGGAAGTCCAGCCTACCAGGCACGTCTGCAGAAACAGTGCAAGGAAGGGCAGCGACTTCCTG
GTTGAGCTTCTGCTAAAACATGGACATGCTTCAGTGCTGCTCCTGAGAGAGTAGCAGGTTAC
CACTCTGGCAGGCCCCAGCCCTGCAGCAAGGAGGAAGAGGACTCAAAAGTCTGGCCTTTCAC
TGAGCCTCCACAGCAGTGGGGGAGAAGCAAGGGTTGGGCCCAGTGTCCCCTTTCCCCAGTGA
CACCTCAGCCTTGGCAGCCCTGATGACTGGTCTCTGGCTGCAACTTAATGCTCTGATATGGC
TTTTAGCATTTATTATATGAAAATAGCAGGGTTTTAGTTTTTAATTTATCAGAGACCCTGCC
ACCCATTCCATCTCCATCCAAGCAAACTGAATGGCAATGAAACAAACTGGAGAAGAAGGTAG
GAGAAAGGGCGGTGAACTCTGGCTCTTTGCTGTGGACATGCGTGACCAGCAGTACTCAGGTT
TGAGGGTTTGCAGAAAGCCAGGGAACCCACAGAGTCACCAACCCTTCATTTAACAAGTAAGA
ATGTTAAAAAGTGAAAACAATGTAAGAGCCTAACTCCATCCCCCGTGGCCATTACTGCATAA
AATAGAGTGCATTTGAAAT

FIGURE 30

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49624

><subunit 1 of 1, 735 aa, 1 stop

><MW: 80177, pI: 7.08, NX(S/T): 5

MAARPLPVSPARALLLALAGALLAPCEARGVSLWNQGRADEVVSASVRSGDLWIPVKSFDSK
NHPEVLNIRLQRESKELIINLERNEGLIASSFTETHYLQDGTDVSLARNYTGHCYYHGHVRG
YSDSAVSLSTCSGLRGLIVFENESYVLEPMKSATNRYKLFPAKKLKSVRGSCGSHHNTPNLA
AKNVFPPPSQTWARRHKRETLKATKYVELVIVADNREFQRQGKDLEKVKQRLIEIANHVDKF
YRPLNIRIVLVGVEVWNDMDKCSVSQDPFTSLHEFLDWRKMKLLPRKSHDNAQLVSGVYFQG
TTIGMAPIMSMCTADQSGGIVMDHSDNPLGAAVTLAHELGHNFGMNHDTLDRGCSCQMAVEK
GGCIMNASTGYPFPMVFSSCSRKDLETSLEKGMGVCLFNLPEVRESFGGQKCGNRFVEEGEE
CDCGEPEECMNRCCNATTCTLKPDAVCAHGLCCEDCQLKPAGTACRDSSNSCDLPEFCTGAS
PHCPANVYLHDGHSCQDVDGYCYNGICQTHEQQCVTLWGPGAKPAPGICFERVNSAGDPYGN
CGKVSKSSFAKCEMRDAKCGKIQCQGGASRPVIGTNAVSIETNIPLQQGGRILCRGTHVYLG
DDMPDPGLVLAGTKCADGKICLNRQCQNISVFGVHECAMQCHGRGVCNNRKNCHCEAHWAPP
FCDKFGFGGSTDSGPIRQAEARQEAAESNRERGQGQEPVGSQEHASTASLTLI

FIGURE 31

TCCCAAGGCTTCTTGGATGGCAGATGATTNTGGGGTTTTGCATTGTTTCCCTGACAACGAAA
ACAAAACAGTTTTGGGGGTTCAGGAGGGGAANTCCAGCCTACCCAGGAAGTTTGCAGAAACA
GTGCAAGGAAGGGCAGGANTTCCTGGTTGAGNTTTTTGNTAAAACATGGACATGNTTCAGTG
CTGCTCNTGAGAGAGTAGCAGGTTACCACTTTTGGCAGGCCCCAGCCCTGCAGCAAGGAGGA
AGAGGACTCAAAAGTTTGGCCTTTCACTGAGCCTCCACAGCAGTGGGGGAGAAGCAAGGGTT
GGGCCCAGTGTCCCCTTTCCCCAGTGACACCTCAGCCTTGGCAGCCCTGATAACTGGTNTNT
GGCTGCAANTTAATGCTNTGATATGGCTTTTAGCATTTATTATATGAAAATAGCAGGGTTTT
AGTTTTTAATTTATCAGAGACCCTGCCACCCATTCCATNTCCATCCAAG

FIGURE 32

CATCCTGCAACATGGTGAAACCACGCCTGGCTAATTTTGTTGTATTTTTGGTAGAGATGGGA
TTTCACCGTGTTAGCCAGGATTGTCTCAATCTGACCTCATGATCTGCCCGCCTCGGCCTCCC
AAAGTGCTGGGATTACAGGCGAGTGCAACCACACCCGGCCACAAACTTTTTAAGAAGTTAAT
GAAACCATACCTTTTACATTTTTAATGACAGGAAAATGCTCACAATAATTGTTAACCCAAAA
TTCTGGATACAAAAGTACAATCTTTACTGTGTAAATACATGTATATGTACTATATGAAAATA
TACCAAATATCAATAATACTTATCTCTGGGTAAAAACCTCTTCTCATACCCTGTGCTAACAA
CTTTTAACAAAAAATTTGCATCACTTTTAAGAATCAAGAAAAATTTCTGAAGGTCATATGGG
ACAGAAAAAAAAACCAAGGGAAAAATCACGCCACTTGGGAAAAAAAGATTCGAAATCTGCCT
TTTTATAGATTTGTAATTAATAAGGTCCAGGCTTTCTAAGCAACTTAAATGTTTTGTTTCGA
AACAAAGTACTTGTCTGGATGTAGGAGGAAAGGGAGTGATGTCACTGCCATTATGATGCCCC
TTGAATATAAGACCCTACTTGCTATCTCCCCTGCACCAGCCAGGAGCCACCCATCCTCCAGC
ACACTGAGCAGCAAGCTGGACACACGGCACACTGATCCAA**ATG**GGTAAGGGGATGGTGGCGA
TGCTCATTCTGGGTCTGCTACTTCTGGCGCTGCTCCTACCCGTGCAGGTTTCTTCATTTGTT
CCTTTAACCAGTATGCCGGAAGCTACTGCAGCCGAAACCACAAAGCCCTCCAACAGTGCCCT
ACAGCCTACAGCCGGTCTCCTTGTGGTCTTGCTTGCCCTTCTACATCTCTACCAT**TAA**GAGG
CAGGTCAAGAAACAGCTACAGTTCTCCAACCCATACACTAAAACCGAATCCAAATGGTGCCT
AGAAGTTCAATGTGGCAAGGAAAAAAACCAGGTCTTCATCAAATCTACTAATTTCACTCCTT
ATTAACAGAGAAACGCTTGAGAGTCTCAAACTGGACTGGTTTAAAGAGCATCTGAAGGATTT
GACTAGATGATAAATGCCTGTACTCCCAGTACTTTGGGAGGCCTAGGCCGGCGGATCACCTG
AGGTCAGGAGTTTGAGACTAACCTGGCCAAAATGGTGAAACCCCATCTGTACTAAAAATACA
AATATTGACTGGGCGTGGTGGTGAGTGCCTGTGATCCCAGCTACTCAGGTGGCTGAAGCAGG
ACAATCACTTGAACTCAGGAGGCAGAGGTTGCAGTGAGCTGAGATCGCGCTACTGCACTCTA
GCCTAGCCTGGGCAACAGAGTGAGACTTCGTCTCAAAAAAAAAAAAGCCAAGTGCAGTGGCT
CACGCCTGTAATCCCGGCACTTTGGGAGGCCGAGGTGGGCGGATCACGAGGTCAGGAGATCA
AGACCATCCTGGCTAATACAGTGAAACCCTGTCTCTACTAAAAATACAAAAAATTAGCCGGG
GATGGTGGCAGGCACCTGGAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATAGCGTGAA
CTCAGGAGGCGGAGCTTGCAGTGAGCCGAGATTGCGCTACTGCACTCCAGCCTGGGCGACAG
CGCGAGACTCCGTCTCAAAAAAAAAAAAAAAAAAAAAA

FIGURE 33

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48309
><subunit 1 of 1, 67 aa, 1 stop
><MW: 6981, pI: 7.47, NX(S/T): 0
MGKGMVAMLILGLLLLALLLPVQVSSFVPLTSMPEATAAETTKPSNSALQPTAGLLVVLLAL
LHLYH

FIGURE 34

GCCGCGGCGAGAGCGCGCCCAGCCCCGCCGCGATGCCCGCGCGCCCAGGACGCCTCCTCCCG
CTGCTGGCCCGGCCGGCGGCCCTGACTGCGCTGCTGCTGCTGCTGCTGGGCCATGGCGGCGG
CGGGCGCTGGGGCGCCCGGGCCCAGGAGGCGGCGGCGGCGGCGGCGGACGGGCCCCCCGCGG
CAGACGGCGAGGACGGACAGGACCCGCACAGCAAGCACCTGTACACGGCCGACATGTTCACG
CACGGGATCCAGAGCGCCGCGCACTTCGTCATGTTCTTCGCGCCCTGGTGTGGACACTGCCA
GCGGCTGCAGCCGACTTGGAATGACCTGGGAGACAAATACAACAGCATGGAAGATGCCAAAG
TCTATGTGGCTAAAGTGGACTGCACGGCCCACTCCGACGTGTGCTCCGCCCAGGGGGTGCGA
GGATACCCCACCTTAAAGCTTTTCAAGCCAGGCCAAGAAGCTGTGAAGTACCAGGGTCCTCG
GGACTTCCAGACACTGGAAAACTGGATGCTGCAGACACTGAACGAGGAGCCAGTGACACCAG
AGCCGGAAGTGGAACCGCCCAGTGCCCCCGAGCTCAAGCAAGGGCTGTATGAGCTCTCAGCA
AGCAACTTTGAGCTGCACGTTGCACAAGGCGACCACTTTATCAAGTTCTTCGCTCCGTGGTG
TGGTCACTGCAAAGCCCTGGCTCCAACCTGGGAGCAGCTGGCTCTGGGCCTTGAACATTCCG
AAACTGTCAAGATTGGCAAGGTTGATTGTACACAGCACTATGAACTCTGCTCCGGAAACCAG
GTTCGTGGCTATCCCACTCTTCTCTGGTTCCGAGATGGGAAAAAGGTGGATCAGTACAAGGG
AAAGCGGGATTTGGAGTCACTGAGGGAGTACGTGGAGTCGCAGCTGCAGCGCACAGAGACTG
GAGCGACGGAGACCGTCACGCCCTCAGAGGCCCCGGTGCTGGCAGCTGAGCCCGAGGCTGAC
AAGGGCACTGTGTTGGCACTCACTGAAAATAACTTCGATGACACCATTGCAGAAGGAATAAC
CTTCATCAAGTTTTATGCTCCATGGTGTGGTCATTGTAAGACTCTGGCTCCTACTTGGGAGG
AACTCTCTAAAAAGGAATTCCCTGGTCTGGCGGGGGTCAAGATCGCCGAAGTAGACTGCACT
GCTGAACGGAATATCTGCAGCAAGTATTCGGTACGAGGCTACCCCACGTTATTGCTTTTCCG
AGGAGGGAAGAAAGTCAGTGAGCACAGTGGAGGCAGAGACCTTGACTCGTTACACCGCTTTG
TCCTGAGCCAAGCGAAAGACGAACTTTAGGAACACAGTTGGAGGTCACCTCTCCTGCCCAGC
TCCCGCACCCTGCGTTTAGGAGTTCAGTCCCACAGAGGCCACTGGGTTCCCAGTGGTGGCTG
TTCAGAAAGCAGAACATACTAAGCGTGAGGTATCTTCTTTGTGTGTGTGTTTTCCAAGCCAA
CACACTCTACAGATTCTTTATTAAGTTAAGTTTCTCTAAGTAAATGTGTAACTCATGGTCAC
TGTGTAAACATTTTCAGTGGCGATATATCCCCTTTGACCTTCTCTTGATGAAATTTACATGG
TTTCCTTTGAGACTAAAATAGCGTTGAGGGAAATGAAATTGCTGGACTATTTGTGGCTCCTG
AGTTGAGTGATTTTGGTGAAAGAAAGCACATCCAAAGCATAGTTTACCTGCCCACGAGTTCT
GGAAAGGTGGCCTTGTGGCAGTATTGACGTTCCTCTGATCTTAAGGTCACAGTTGACTCAAT
ACTGTGTTGGTCCGTAGCATGGAGCAGATTGAAATGCAAAAACCCACACCTCTGGAAGATAC
CTTCACGGCCGCTGCTGGAGCTTCTGTTGCTGTGAATACTTCTCTCAGTGTGAGAGGTTAGC
CGTGATGAAAGCAGCGTTACTTCTGACCGTGCCTGAGTAAGAGAATGCTGATGCCATAACTT
TATGTGTCGATACTTGTCAAATCAGTTACTGTTCAGGGGATCCTTCTGTTTCTCACGGGGTG
AAACATGTCTTTAGTTCCTCATGTTAACACGAAGCCAGAGCCCACATGAACTGTTGGATGTC
TTCCTTAGAAAGGGTAGGCATGGAAAATTCCACGAGGCTCATTCTCAGTATCTCATTAACTC
ATTGAAAGATTCCAGTTGTATTTGTCACCTGGGGTGACAAGACCAGACAGGCTTTCCCAGGC
CTGGGTATCCAGGGAGGCTCTGCAGCCCTGCTGAAGGGCCCTAACTAGAGTTCTAGAGTTTC
TGATTCTGTTTCTCAGTAGTCCTTTTAGAGGCTTGCTATACTTGGTCTGCTTCAAGGAGGTC
GACCTTCTAATGTATGAAGAATGGGATGCATTTGATCTCAAGACCAAAGACAGATGTCAGTG
GGCTGCTCTGGCCCTGGTGTGCACGGCTGTGGCAGCTGTTGATGCCAGTGTCCTCTAACTCA
TGCTGTCCTTGTGATTAAACACCTCTATCTCCCTTGGGAATAAGCACATACAGGCTTAAGCT
CTAAGATAGATAGGTGTTTGTCCTTTTACCATCGAGCTACTTCCCATAATAACCACTTTGCA
TCCAACACTCTTCACCCACCTCCCATACGCAAGGGGATGTGGATACTTGGCCCAAAGTAACT
GGTGGTAGGAATCTTAGAAACAAGACCACTTATACTGTCTGTCTGAGGCAGAAGATAACAGC
AGCATCTCGACCAGCCTCTGCCTTAAAGGAAATCTTTATTAATCACGTATGGTTCACAGATA
ATTCTTTTTTTTAAAAAAACCCAACCTCCTAGAGAAGCACAACTGTCAAGAGTCTTGTACACA
CAACTTCAGCTTTGCATCACGAGTCTTGTATTCCAAGAAAATCAAAGTGGTACAATTTGTTT
GTTTACACTATGATACTTTCTAAATAAACTCTTTTTTTTTAA

FIGURE 35

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA46776
><subunit 1 of 1, 432 aa, 1 stop
><MW: 47629, pI: 5.90, NX(S/T): 0

MPARPGRLLPLLARPAALTALLLLLLGHGGGGRWGARAQEAAAAAADGPPAADGEDGQDPHS
KHLYTADMFTHGIQSAAHFVMFFAPWCGHCQRLQPTWNDLGDKYNSMEDAKVYVAKVDCTAH
SDVCSAQGVRGYPTLKLFKPGQEAVKYQGPRDFQTLENWMLQTLNEEPVTPEPEVEPPSAPE
LKQGLYELSASNFELHVAQGDHFIKFFAPWCGHCKALAPTWEQLALGLEHSETVKIGKVDCT
QHYELCSGNQVRGYPTLLWFRDGKKVDQYKGKRDLESLREYVESQLQRTETGATETVTPSEA
PVLAAEPEADKGTVLALTENNFDDTIAEGITFIKFYAPWCGHCKTLAPTWEELSKKEFPGLA
GVKIAEVDCTAERNICSKYSVRGYPTLLLFRGGKKVSEHSGGRDLDSLHRFVLSQAKDEL

FIGURE 36

CTTTTCTGAGGAACCACAGCA**ATG**AATGGCTTTGCATCCTTGCTTCGAAGAAACCAATTTAT
CCTCCTGGTACTATTTCTTTTGCAAATTCAGAGTCTGGGTCTGGATATTGATAGCCGTCCTA
CCGCTGAAGTCTGTGCCACACACACAATTTCACCAGGACCCAAAGGAGATGATGGTGAAAAA
GGAGATCCAGGAGAAGAGGGAAAGCATGGCAAAGTGGGACGCATGGGGCCGAAAGGAATTAA
AGGAGAACTGGGTGATATGGGAGATCAGGGCAATATTGGCAAGACTGGGCCCATTGGGAAGA
AGGGTGACAAAGGGGAAAAAGGTTTGCTTGGAATACCTGGAGAAAAAGGCAAAGCAGGTACT
GTCTGTGATTGTGGAAGATACCGGAAATTTGTTGGACAACTGGATATTAGTATTGCTCGGCT
CAAGACATCTATGAAGTTTGTCAAGAATGTGATAGCAGGGATTAGGGAAACTGAAGAGAAAT
TCTACTACATCGTGCAGGAAGAGAAGAACTACAGGGAATCCCTAACCCACTGCAGGATTCGG
GGTGGAATGCTAGCCATGCCCAAGGATGAAGCTGCCAACACACTCATCGCTGACTATGTTGC
CAAGAGTGGCTTCTTTCGGGTGTTCATTGGCGTGAATGACCTTGAAAGGGAGGGACAGTACA
TGTCCACAGACAACACTCCACTGCAGAACTATAGCAACTGGAATGAGGGGGAACCCAGCGAC
CCCTATGGTCATGAGGACTGTGTGGAGATGCTGAGCTCTGGCAGATGGAATGACACAGAGTG
CCATCTTACCATGTACTTTGTCTGTGAGTTCATCAAGAAGAAAAAG**TAA**CTTCCCTCATCCT
ACGTATTTGCTATTTTCCTGTGACCGTCATTACAGTTATTGTTATCCATCCTTTTTTTCCTG
ATTGTACTACATTTGATCTGAGTCAACATAGCTAGAAAATGCTAAACTGAGGTATGGAGCCT
CCATCATCAAAAAAAAAAAAAAAA

FIGURE 37

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50980
><subunit 1 of 1, 277 aa, 1 stop
><MW: 30645, pI: 7.47, NX(S/T): 2
MNGFASLLRRNQFILLVLFLLQIQSLGLDIDSRPTAEVCATHTISPGPKGDDGEKGDPGEEG
KHGKVGRMGPKGIKGELGDMGDQGNIGKTGPIGKKGDKGEKGLLGIPGEKGKAGTVCDCGRY
RKFVGQLDISIARLKTSMKFVKNVIAGIRETEEKFYYIVQEEKNYRESLTHCRIRGGMLAMP
KDEAANTLIADYVAKSGFFRVFIGVNDLEREGQYMSTDNTPLQNYSNWNEGEPSDPYGHEDC
VEMLSSGRWNDTECHLTMYFVCEFIKKKK

FIGURE 38

GGTTCTATCGATTCGAATTCGGCCACACTGGCCGGATCCTCTAGAGATCCCTCGACCTCGAC
CCACGCGTCCGCTGCTCTCCGCCCGTGTGGAGTGGTGGGGGCCTGGGTGGGAATGGGCGTGT
GCCAGCGCACGCGCGCTCCCTGGAAGGAGAAGTCTCAGCTAGAACGAGCGGCCCTAGGTTTT
CGGAAGGGAGGATCAGGGATGTTTGCGAGCGGCTGGAACCAGACGGTGCCGATAGAGGAAGC
GGGCTCCATGGCTGCCCTCCTGCTGCTGCCCCTGCTGCTGTTGCTACCGCTGCTGCTGCTGA
AGCTACACCTCTGGCCGCAGTTGCGCTGGCTTCCGGCGGACTTGGCCTTTGCGGTGCGAGCT
CTGTGCTGCAAAAGGGCTCTTCGAGCTCGCGCCCTGGCCGCGGCTGCCGCCGACCCGGAAGG
TCCCGAGGGGGGCTGCAGCCTGGCCTGGCGCCTCGCGGAACTGGCCCAGCAGCGCGCCGCGC
ACACCTTTCTCATTCACGGCTCGCGGCGCTTTAGCTACTCAGAGGCGGAGCGCGAGAGTAAC
AGGGCTGCACGCGCCTTCCTACGTGCGCTAGGCTGGGACTGGGGACCCGACGGCGGCGACAG
CGGCGAGGGGAGCGCTGGAGAAGGCGAGCGGGCAGCGCCGGGAGCCGGAGATGCAGCGGCCG
GAAGCGGCGCGGAGTTTGCCGGAGGGGACGGTGCCGCCAGAGGTGGAGGAGCCGCCGCCCCT
CTGTCACCTGGAGCAACTGTGGCGCTGCTCCTCCCCGCTGGCCCAGAGTTTCTGTGGCTCTG
GTTCGGGCTGGCCAAGGCCGGCCTGCGCACTGCCTTTGTGCCCACCGCCCTGCGCCGGGGCC
CCCTGCTGCACTGCCTCCGCAGCTGCGGCGCGCGCGCGCTGGTGCTGGCGCCAGAGTTTCTG
GAGTCCCTGGAGCCGGACCTGCCCGCCCTGAGAGCCATGGGGCTCCACCTGTGGGCTGCAGG
CCCAGGAACCCACCCTGCTGGAATTAGCGATTTGCTGGCTGAAGTGTCCGCTGAAGTGGATG
GGCCAGTGCCAGGATACCTCTCTTCCCCCCAGAGCATAACAGACACGTGCCTGTACATCTTC
ACCTCTGGCACCACGGGCCTCCCCAAGGCTGCTCGGATCAGTCATCTGAAGATCCTGCAATG
CCAGGGCTTCTATCAGCTGTGTGGTGTCCACCAGGAAGATGTGATCTACCTCGCCCTCCCAC
TCTACCACATGTCCGGTTCCCTGCTGGGCATCGTGGGCTGCATGGGCATTGGGGCCACAGTG
GTGCTGAAATCCAAGTTCTCGGCTGGTCAGTTCTGGGAAGATTGCCAGCAGCACAGGGTGAC
GGTGTTCCAGTACATTGGGGAGCTGTGCCGATACCTTGTCAACCAGCCCCCGAGCAAGGCAG
AACGTGGCCATAAGGTCCGGCTGGCAGTGGGCAGCGGGCTGCGCCCAGATACCTGGGAGCGT
TTTGTGCGGCGCTTCGGGCCCCTGCAGGTGCTGGAGACATATGGACTGACAGAGGGCAACGT
GGCCACCATCAACTACACAGGACAGCGGGGCGCTGTGGGGCGTGCTTCCTGGCTTTACAAGC
ATATCTTCCCCTTCTCCTTGATTCGCTATGATGTCACCACAGGAGAGCCAATTCGGGACCCC
CAGGGGCACTGTATGGCCACATCTCCAGGTGAGCCAGGGCTGCTGGTGGCCCCGGTAAGCCA
GCAGTCCCCATTCCTGGGCTATGCTGGCGGGCCAGAGCTGGCCCAGGGGAAGTTGCTAAAGG
ATGTCTTCCGGCCTGGGGATGTTTTCTTCAACACTGGGGACCTGCTGGTCTGCGATGACCAA
GGTTTTCTCCGCTTCCATGATCGTACTGGAGACACCTTCAGGTGGAAGGGGGAGAATGTGGC
CACAACCGAGGTGGCAGAGGTCTTCGAGGCCCTAGATTTTCTTCAGGAGGTGAACGTCTATG
GAGTCACTGTGCCAGGGCATGAAGGCAGGGCTGGAATGGCAGCCCTAGTTCTGCGTCCCCCC
CACGCTTTGGACCTTATGCAGCTCTACACCCACGTGTCTGAGAACTTGCCACCTTATGCCCG
GCCCCGATTCCTCAGGCTCCAGGAGTCTTTGGCCACCACAGAGACCTTCAAACAGCAGAAAG
TTCGGATGGCAAATGAGGGCTTCGACCCCAGCACCCTGTCTGACCCACTGTACGTTCTGGAC
CAGGCTGTAGGTGCCTACCTGCCCCTCACAACTGCCCGGTACAGCGCCCTCCTGGCAGGAAA
CCTTCGAATCTGAGAACTTCCACACCTGAGGCACCTGAGAGAGGAACTCTGTGGGGTGGGGG
CCGTTGCAGGTGTACTGGGCTGTCAGGGATCTTTTCTATACCAGAACTGCGGTCACTATTTT
GTAATAAATGTGGCTGGAGCTGATCCAGCTGTCTCTGACCTAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAGGGCGGCCGCGACTCTAGAGTCGACCTGCAGTAGGGATAACAGGGTAATAAGC
TTGGCCGCCATGGCCCAACTTGTTTATTGCAG

FIGURE 39

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50913
><subunit 1 of 1, 730 aa, 1 stop
><MW: 78644, pI: 7.65, NX(S/T): 2

MGVCQRTRAPWKEKSQLERAALGFRKGGSGMFASGWNQTVPIEEAGSMAALLLLPLLLLLPL
LLLKLHLWPQLRWLPADLAFAVRALCCKRALRARALAAAAADPEGPEGGCSLAWRLAELAQQ
RAAHTFLIHGSRRFSYSEAERESNRAARAFLRALGWDWGPDGGDSGEGSAGEGERAAPGAGD
AAAGSGAEFAGGDGAARGGGAAAPLSPGATVALLLPAGPEFLWLWFGLAKAGLRTAFVPTAL
RRGPLLHCLRSCGARALVLAPEFLESLEPDLPALRAMGLHLWAAGPGTHPAGISDLLAEVSA
EVDGPVPGYLSSPQSITDTCLYIFTSGTTGLPKAARISHLKILQCQGFYQLCGVHQEDVIYL
ALPLYHMSGSLLGIVGCMGIGATVVLKSKFSAGQFWEDCQQHRVTVFQYIGELCRYLVNQPP
SKAERGHKVRLAVGSGLRPDTWERFVRRFGPLQVLETYGLTEGNVATINYTGQRGAVGRASW
LYKHIFPFSLIRYDVTTGEPIRDPQGHCMATSPGEPGLLVAPVSQQSPFLGYAGGPELAQGK
LLKDVFRPGDVFFNTGDLLVCDDQGFLRFHDRTGDTFRWKGENVATTEVAEVFEALDFLQEV
NVYGVTVPGHEGRAGMAALVLRPPHALDLMQLYTHVSENLPPYARPRFLRLQESLATTETFK
QQKVRMANEGFDPSTLSDPLYVLDQAVGAYLPLTTARYSALLAGNLRI

Signal peptide:
aa 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation site
starting at aa 136

CUB domain protein motif
aa 254-261

putative AMP-binding domain siganture
aa 332-343

N-glycosylation sites
aa 37-40 and 483-486

FIGURE 40

CCTGTGTTAAGCTGAGGTTTCCCCTAGATCTCGTATATCCCCAACACATACCTCCACGCACA
CACATCCCCAAGAACCTCGAGCTCACACCAACAGACACACGCGCGCATACACACTCGCTCTC
GCTTGTCCATCTCCCTCCCGGGGGAGCCGGCGCGCGCTCCCACCTTTGCCGCACACTCCGGC
GAGCCGAGCCCGCAGCGCTCCAGGATTCTGCGGCTCGGAACTCGGATTGCAGCTCTGAACCC
CCATGGTGGTTTTTTAAACACTTCTTTTCCTTCTCTTCCTCGTTTTGATTGCACCGTTTCCA
TCTGGGGGCTAGAGGAGCAAGGCAGCAGCCTTCCCAGCCAGCCCTTGTTGGCTTGCCATCGT
CCATCTGGCTTATAAAAGTTTGCTGAGCGCAGTCCAGAGGGCTGCGCTGCTCGTCCCCTCGG
CTGGCAGAAGGGGGTGACGCTGGGCAGCGGCGAGGAGCGCGCCGCTGCCTCTGGCGGGCTTT
CGGCTTGAGGGGCAAGGTGAAGAGCGCACCGGCCGTGGGGTTTACCGAGCTGGATTTGTATG
TTGCACC**ATG**CCTTCTTGGATCGGGGCTGTGATTCTTCCCCTCTTGGGGCTGCTGCTCTCCC
TCCCCGCCGGGGCGGATGTGAAGGCTCGGAGCTGCGGAGAGGTCCGCCAGGCGTACGGTGCC
AAGGGATTCAGCCTGGCGGACATCCCCTACCAGGAGATCGCAGGGGAACACTTAAGAATCTG
TCCTCAGGAATATACATGCTGCACCACAGAAATGGAAGACAAGTTAAGCCAACAAAGCAAAC
TCGAATTTGAAAACCTTGTGGAAGAGACAAGCCATTTTGTGCGCACCACTTTTGTGTCCAGG
CATAAGAAATTTGACGAATTTTTCCGAGAGCTCCTGGAGAATGCAGAAAAGTCACTAAATGA
TATGTTTGTACGGACCTATGGCATGCTGTACATGCAGAATTCAGAAGTCTTCCAGGACCTCT
TCACAGAGCTGAAAAGGTACTACACTGGGGGTAATGTGAATCTGGAGGAAATGCTCAATGAC
TTTTGGGCTCGGCTCCTGGAACGGATGTTTCAGCTGATAAACCCTCAGTATCACTTCAGTGA
AGACTACCTGGAATGTGTGAGCAAATACACTGACCAGCTCAAGCCATTTGGAGACGTGCCCC
GGAAACTGAAGATTCAGGTTACCCGCGCCTTCATTGCTGCCAGGACCTTTGTCCAGGGGCTG
ACTGTGGGCAGAGAAGTTGCAAACCGAGTTTCCAAGGTCAGCCCAACCCCAGGGTGTATCCG
TGCCCTCATGAAGATGCTGTACTGCCCATACTGTCGGGGGCTTCCCACTGTGAGGCCCTGCA
ACAACTACTGTCTCAACGTCATGAAGGGCTGCTTGGCAAATCAGGCTGACCTCGACACAGAG
TGGAATCTGTTTATAGATGCAATGCTCTTGGTGGCAGAGCGACTGGAGGGGCCATTCAACAT
TGAGTCGGTCATGGACCCGATAGATGTCAAGATTTCTGAAGCCATTATGAACATGCAAGAAA
ACAGCATGCAGGTGTCTGCAAAGGTCTTTCAGGGATGTGGTCAGCCCAAACCTGCTCCAGCC
CTCAGATCTGCCCGCTCAGCTCCTGAAAATTTTAATACACGTTTCAGGCCCTACAATCCTGA
GGAAAGACCAACAACTGCTGCAGGCACAAGCTTGGACCGGCTGGTCACAGACATAAAAGAGA
AATTGAAGCTCTCTAAAAAGGTCTGGTCAGCATTACCCTACACTATCTGCAAGGACGAGAGC
GTGACAGCGGGCACGTCCAACGAGGAGGAATGCTGGAACGGGCACAGCAAAGCCAGATACTT
GCCTGAGATCATGAATGATGGGCTCACCAACCAGATCAACAATCCCGAGGTGGATGTGGACA
TCACTCGGCCTGACACTTTCATCAGACAGCAGATTATGGCTCTCCGTGTGATGACCAACAAA
CTAAAAAACGCCTACAATGGCAATGATGTCAATTTCCAGGACACAAGTGATGAATCCAGTGG
CTCAGGGAGTGGCAGTGGGTGCATGGATGACGTGTGTCCCACGGAGTTTGAGTTTGTCACCA
CAGAGGCCCCCGCAGTGGATCCCGACCGGAGAGAGGTGGACTCTTCTGCAGCCCAGCGTGGC
CACTCCCTGCTCTCCTGGTCTCTCACCTGCATTGTCCTGGCACTGCAGAGACTGTGCAGA**TA**
ATCTTGGGTTTTTGGTCAGATGAAACTGCATTTTAGCTATCTGAATGGCCAACTCACTTCTT
TTCTTACACTCTTGGACAATGGACCATGCCACAAAAACTTACCGTTTTCTATGAGAAGAGAG
CAGTAATGCAATCTGCCTCCCTTTTTGTTTTCCCAAAGAGTACCGGGTGCCAGACTGAACTG
CTTCCTCTTTCCTTCAGCTATCTGTGGGGACCTTGTTTATTCTAGAGAGAATTCTTACTCAA
ATTTTTCGTACCAGGAGATTTTCTTACCTTCATTTGCTTTTATGCTGCAGAAGTAAAGGAAT
CTCACGTTGTGAGGGTTTTTTTTTCTCATTTAAAAT

FIGURE 41

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50914
><subunit 1 of 1, 555 aa, 1 stop
><MW: 62736, pI: 5.36, NX(S/T): 0
MPSWIGAVILPLLGLLLSLPAGADVKARSCGEVRQAYGAKGFSLADIPYQEIAGEHLRICPQ
EYTCCTTEMEDKLSQQSKLEFENLVEETSHFVRTTFVSRHKKFDEFFRELLENAEKSLNDMF
VRTYGMLYMQNSEVFQDLFTELKRYYTGGNVNLEEMLNDFWARLLERMFQLINPQYHFSEDY
LECVSKYTDQLKPFGDVPRKLKIQVTRAFIAARTFVQGLTVGREVANRVSKVSPTPGCIRAL
MKMLYCPYCRGLPTVRPCNNYCLNVMKGCLANQADLDTEWNLFIDAMLLVAERLEGPFNIES
VMDPIDVKISEAIMNMQENSMQVSAKVFQGCGQPKPAPALRSARSAPENFNTRFRPYNPEER
PTTAAGTSLDRLVTDIKEKLKLSKKVWSALPYTICKDESVTAGTSNEEECWNGHSKARYLPE
IMNDGLTNQINNPEVDVDITRPDTFIRQQIMALRVMTNKLKNAYNGNDVNFQDTSDESSGSG
SGSGCMDDVCPTEFEFVTTEAPAVDPDRREVDSSAAQRGHSLLSWSLTCIVLALQRLCR

FIGURE 42A

CGGACGCGTGGGCGGACGCGTGGGCAAAAGAACTCGGAGTGCCAAAGCTAAATAAGTTAGCT
GAGAAAACGCACGCAGTTTGCAGCGCCTGCGCCGGGTGCGCCAACTACGCAAAGACCAAGCG
GGCTCCGCGCGGACCGGCCGCGGGGCTAGGGACCCGGCTTTGGCCTTCAGGCTCCCTAGCAG
CGGGGAAAAGGAATTGCTGCCCGGAGTTTCTGCGGAGGTGGAGGGAGATCAGGAAACGGCTT
CTTCCTCACTTCGCCGCCTGGTGAGTGTCGGGGAGATTGGCAAACGCCTAGGAAAGGACTGG
GGAAAATAGCCCTGGGAAAGTGGAGAAGGTGATCAGGAGGCCGGTCCACTACGGCAGTTTAT
CTGTCTGATCAGAGCCAGACGCGACGCGTCCACTTCGCAGTTCTTTCCAGGTGTGGGGACCG
CAGGACAGACGGCCGATCCCGCCGCCCTCCGTACCAGCACTCCCAGGAGAGTCAGCCTCGCT
CCCCAACGTCGAGGGCGCTCTGGCCACGAAAAGTTCCTGTCCACTGTGATTCTCAATTCCTT
GCTTGGTTTTTTTCTCCAGAGAACTTTTGGGTGGAGATATTAACTTTTTTCTTTTTTTTTTT
CCTTGGTGGAAGCTGCTCTAGGGAGGGGGGAGGAGGAGGAGAAAGTGAAATGTGCTGGAGAA
GAGCGAGCCCTCCTTGTTCTTCCGGAGTCCCATCCATTAAGCCATCACTTCTGGAAGATTAA
AGTTGTCGGACATGGTGACAGCTGAGAGGAGAGGAGGATTTCTTGCCAGGTGGAGAGTCTTC
ACCGTCTGTTGGGTGCATGTGTGCGCCCGCAGCGGCGCGGGGCGCGTGGTTCTCCGCGTGGA
GTCTCACCTGGGACCTGAGTGA**ATG**GCTCCCAGGGGCTGTGCGGGGCATCCGCCTCCGCCTT
CTCCACAGGCCTGTGTCTGTCCTGGAAAGATGCTAGCAATGGGGGCGCTGGCAGGATTCTGG
ATCCTCTGCCTCCTCACTTATGGTTACCTGTCCTGGGGCCAGGCCTTAGAAGAGGAGGAAGA
AGGGGCCTTACTAGCTCAAGCTGGAGAGAAACTAGAGCCCAGCACAACTTCCACCTCCCAGC
CCCATCTCATTTTCATCCTAGCGGATGATCAGGGATTTAGAGATGTGGGTTACCACGGATCT
GAGATTAAAACACCTACTCTTGACAAGCTCGCTGCCGAAGGAGTTAAACTGGAGAACTACTA
TGTCCAGCCTATTTGCACACCATCCAGGAGTCAGTTTATTACTGGAAAGTATCAGATACACA
CCGGACTTCAACATTCTATCATAAGACCTACCCAACCCAACTGTTTACCTCTGGACAATGCC
ACCCTACCTCAGAAACTGAAGGAGGTTGGATATTCAACGCATATGGTCGGAAAATGGCACTT
GGGTTTTAACAGAAAAGAATGCATGCCCACCAGAAGAGGATTTGATACCTTTTTTGGTTCCC
TTTTGGGAAGTGGGGATTACTATACACACTACAAATGTGACAGTCCTGGGATGTGTGGCTAT
GACTTGTATGAAAACGACAATGCTGCCTGGGACTATGACAATGGCATATACTCCACACAGAT
GTACACTCAGAGAGTACAGCAAATCTTAGCTTCCCATAACCCCACAAAGCCTATATTTTTAT
ATACTGCCTATCAAGCTGTTCATTCACCACTGCAAGCTCCTGGCAGGTATTTCGAACACTAC
CGATCCATTATCAACATAAACAGGAGAAGATATGCTGCCATGCTTTCCTGCTTAGATGAAGC
AATCAACAACGTGACATTGGCTCTAAAGACTTATGGTTTCTATAACAACAGCATTATCATTT
ACTCTTCAGATAATGGTGGCCAGCCTACGGCAGGAGGGAGTAACTGGCCTCTCAGAGGTAGC
AAAGGAACATATTGGGAAGGAGGGATCCGGGCTGTAGGCTTTGTGCATAGCCCACTTCTGAA
AAACAAGGGAACAGTGTGTAAGGAACTTGTGCACATCACTGACTGGTACCCCACTCTCATTT
CACTGGCTGAAGGACAGATTGATGAGGACATTCAACTAGATGGCTATGATATCTGGGAGACC
ATAAGTGAGGGTCTTCGCTCACCCCGAGTAGATATTTTGCATAACATTGACCCCTATACACC
AAGGCAAAAAATGGCTCCTGGGCAGCAGGCTATGGGATCTGGAACACTGCAATCCAGTCAGC
CATCAGAGTGCAGCACTGGAAATTGCTTACAGGAAATCCTGGCTACAGCGACTGGGTCCCCC
CTCAGTCTTTCAGCAACCTGGGACCGAACCGGTGGCACAATGAACGGATCACCTTGTCAACT
GGCAAAAGTGTATGGCTTTTCAACATCACAGCCGACCCATATGAGAGGGTGGACCTATCTAA
CAGGTATCCAGGAATCG**TGA**AGAAGCTCCTACGGAGGCTCTCACAGTTCAACAAAACTGCAG
TGCCGGTCAGGTATCCCCCCAAAGACCCCAGAAGTAACCCTAGGCTCAATGGAGGGGTCTGG
GGACCATGGTATAAAGAGGAAACCAAGAAAAAGAAGCCAAGCAAAAATCAGGCTGAGAAAAA
GCAAAAGAAAAGCAAAAAAAAGAAGAAGAAACAGCAGAAAGCAGTCTCAGGTAAACCAGCAA
ATTTGGCTCGATAATATCGCTGGCCTAAGCGTCAGGCTTGTTTTCATGCTGTGCCACTCCAG
AGACTTCTGCCACCTGGCCGCCACACTGAAAACTGTCCTGCTCAGTGCCAAGGTGCTACTCT
TGCAAGCCACACTTAGAGAGAGTGGAGATGTTTATTTCTCTCGCTCCTTTAGAAAACGTGGT
GAGTCCTGAGTTCCACTGCTGTGCTTCAGTCAACTGACCAAACACTGCTTTGAATTATAGGA
GGAGAACAATAACCTACCATCCGCAAGCATGCTAATTTGATGGAAGTTACAGGGTAGCATGA
TTAAAACTACCTTTGATAAATTACAGTCAAAGATTGTGTCACCTCAAAGGCCTTGAAGAATA
TATTTTCTTGGTGAATTTTTGTATGTCTGTCATATGACACTTGGGTTTTTTAATTAATTCTA
TTTTATATATATAAATATATGTTTCTTTTCCTGTGAAAAGCTGTTTTTCTCACATGTGAACA
GCTTGCACCTCATTTTACCATGCGTGAGGGAATGGCAAATAAGAATGTTTGAGCACACTGCC
CACAATGAATGTAACTATTTTCTAAACACTTTACTAGAAGAACATTTCAGTATAAAAAACCT
AATTTATTTTTACAGAAAAATATTTTGTTGTTTTTATAAAAAGTTATGCAAATGACTTTTAT
TTTTATTTCCTGCATACCATTAGAAGAATTTTATTTCATTTCTTCAAATTATCAAGCACTGT
AATACTATAAATTAATGTAATACTGTGTGAATTCAGACTATAAAAAACATCATTCAGAAAAC
TTTATAATCGTCATTGTTCAATCAAGATTTTGAATGTAATAAGATGAATATATTCCTTACAA

FIGURE 42B

ATTACTTGGAAATTCAATGTTTGTGCAGAGTTGAGACAACTTTATTGTTTCTATCATAAACT
ATTTATGTATCTTAATTATTAAAATGATTTACTTTATGGCACTAGAAAATTTACTGTGGCTT
TTCTGATCTAACTTCTAGCTAAAATTGTATCATTGGTCCTAAAAAATAAAAATCTTTACTAA
TAGGCAATTGAAGGAATGGTTTGCTAACAACCACAGTAATATAATATGATTTTACAGATAGA
TGCTTCCCCTTGGCTATGACATGGAGAAAGATTTTCCCATAATAATAACTAATATTTATATT
AGGTTGGTGCAAAACTAGTTGCGGTTTTTCCCATTAAAAGTAATAACCTTACTCTTATACAA
AGTGGACACTGTGGGGAGATACAGAGAAATGGAAGATACGGATCCTGCCTGGAGTAGGTAAC
CTTGCTTGGAAACCCCACATGCAAACGTCATGAGGAGAATTAAAGGAGTATTATCAGTAATG
AAGTTTATCATGGGTCATCAATGAGCATAGATTGGTGTGGATCCTGTAGACCCTGGTGTTTT
CTTTGAAGTGCCCTCTCCTAATGCAGAGGCCTTGAAGCTTACAGTATACACTTGAAAAGTCA
CAGATAGCTAGAATTATGATCTTTGAAGTTATAACTGTGATCTGAAAATGTGTGTGGTGGTA
TGACAGCATACCATTAAATACATTTACATCACAGCTCAAAGGACTGTGATATAATCCATTTA
TATCACAACTCAAAGGACTGTGATATAATCCATTTATATCACAGCTCACAGTTTCTGAAAAT
GTATAAAAGAATCTATAATCTAGTACTGAAATTACTAAATTGGGTAAGATGATTTAAATGAT
TTTAATTTTAACATTTTATTTCTAGAATATATGGCTCCATTTTATTTTATAGTGTAAAGTTG
TATTTCCTAAAGTTTGTGTTTTGTCGACAGTATCTTTTAAATGAGTCTTAAAAATAAAGGCA
TATTGTTCATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 43

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48296
><subunit 1 of 1, 515 aa, 1 stop
><MW: 56885, pI: 6.49, NX(S/T): 5

MAPRGCAGHPPPPSPQACVCPGKMLAMGALAGFWILCLLTYGYLSWGQALEEEEEGALLAQA
GEKLEPSTTSTSQPHLIFILADDQGFRDVGYHGSEIKTPTLDKLAAEGVKLENYYVQPICTP
SRSQFITGKYQIHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHLGFNRKEC
MPTRRGFDTFFGSLLGSGDYYTHYKCDSPGMCGYDLYENDNAAWDYDNGIYSTQMYTQRVQQ
ILASHNPTKPIFLYTAYQAVHSPLQAPGRYFEHYRSIININRRRYAAMLSCLDEAINNVTLA
LKTYGFYNNSIIIYSSDNGGQPTAGGSNWPLRGSKGTYWEGGIRAVGFVHSPLLKNKGTVCK
ELVHITDWYPTLISLAEGQIDEDIQLDGYDIWETISEGLRSPRVDILHNIDPYTPRQKMAPG
QQAMGSGTLQSSQPSECSTGNCLQEILATATGSPLSLSATWDRTGGTMNGSPCQLAKVYGFS
TSQPTHMRGWTYLTGIQES

Important Features:

Signal Peptide:
amino acids 1-37

Sulfatases signature 1.
amino acids 120-132

Sulfatases signature 2.
amino acids 168-177

Tyrosine kinase phosphorylation site.
amino acids 163-169

N-glycosylation sites.
amino acids 157-160, 306-309 and 318-321

FIGURE 44

CGGACGCGTGGGTGCGAGTGGAGCGGAGGACCCGAGCGGCTGAGGAGAGAGGAGGCGGCGGC
TTAGCTGCTACGGGGTCCGGCCGGCGCCCTCCCGAGGGGGGCTCAGGAGGAGGAAGGAGGAC
CCGTGCGAGA**ATG**CCTCTGCCCTGGAGCCTTGCGCTCCCGCTGCTGCTCTCCTGGGTGGCAG
GTGGTTTCGGGAACGCGGCCAGTGCAAGGCATCACGGGTTGTTAGCATCGGCACGTCAGCCT
GGGGTCTGTCACTATGGAACTAAACTGGCCTGCTGCTACGGCTGGAGAAGAAACAGCAAGGG
AGTCTGTGAAGCTACATGCGAACCTGGATGTAAGTTTGGTGAGTGCGTGGGACCAAACAAAT
GCAGATGCTTTCCAGGATACACCGGGAAAACCTGCAGTCAAGATGTGAATGAGTGTGGAATG
AAACCCCGGCCATGCCAACACAGATGTGTGAATACACACGGAAGCTACAAGTGCTTTTGCCT
CAGTGGCCACATGCTCATGCCAGATGCTACGTGTGTGAACTCTAGGACATGTGCCATGATAA
ACTGTCAGTACAGCTGTGAAGACACAGAAGAAGGGCCACAGTGCCTGTGTCCATCCTCAGGA
CTCCGCCTGGCCCCAAATGGAAGAGACTGTCTAGATATTGATGAATGTGCCTCTGGTAAAGT
CATCTGTCCCTACAATCGAAGATGTGTGAACACATTTGGAAGCTACTACTGCAAATGTCACA
TTGGTTTCGAACTGCAATATATCAGTGGACGATATGACTGTATAGATATAAATGAATGTACT
ATGGATAGCCATACGTGCAGCCACCATGCCAATTGCTTCAATACCCAAGGGTCCTTCAAGTG
TAAATGCAAGCAGGGATATAAAGGCAATGGACTTCGGTGTTCTGCTATCCCTGAAAATTCTG
TGAAGGAAGTCCTCAGAGCACCTGGTACCATCAAAGACAGAATCAAGAAGTTGCTTGCTCAC
AAAAACAGCATGAAAAAGAAGGCAAAAATTAAAAATGTTACCCCAGAACCCACCAGGACTCC
TACCCCTAAGGTGAACTTGCAGCCCTTCAACTATGAAGAGATAGTTTCCAGAGGCGGGAACT
CTCATGGAGGTAAAAAAGGGAATGAAGAGAAA**TGA**AAGAGGGGCTTGAGGATGAGAAAAGAG
AAGAGAAAGCCCTGAAGAATGACATAGAGGAGCGAAGCCTGCGAGGAGATGTGTTTTTCCCT
AAGGTGAATGAAGCAGGTGAATTCGGCCTGATTCTGGTCCAAAGGAAAGCGCTAACTTCCAA
ACTGGAACATAAAGATTTAAATATCTCGGTTGACTGCAGCTTCAATCATGGGATCTGTGACT
GGAAACAGGATAGAGAAGATGATTTTGACTGGAATCCTGCTGATCGAGATAATGCTATTGGC
TTCTATATGGCAGTTCCGGCCTTGGCAGGTCACAAGAAAGACATTGGCCGATTGAAACTTCT
CCTACCTGACCTGCAACCCCAAAGCAACTTCTGTTTGCTCTTTGATTACCGGCTGGCCGGAG
ACAAAGTCGGGAAACTTCGAGTGTTTGTGAAAAACAGTAACAATGCCCTGGCATGGGAGAAG
ACCACGAGTGAGGATGAAAAGTGGAAGACAGGGAAAATTCAGTTGTATCAAGGAACTGATGC
TACCAAAAGCATCATTTTTGAAGCAGAACGTGGCAAGGGCAAAACCGGCGAAATCGCAGTGG
ATGGCGTCTTGCTTGTTTCAGGCTTATGTCCAGATAGCCTTTTATCTGTGGATGACTGAATG
TTACTATCTTTATATTTGACTTTGTATGTCAGTTCCCTGGTTTTTTTGATATTGCATCATAG
GACCTCTGGCATTTTAGAATTACTAGCTGAAAAATTGTAATGTACCAACAGAAATATTATTG
TAAGATGCCTTTCTTGTATAAGATATGCCAATATTTGCTTTAAATATCATATCACTGTATCT
TCTCAGTCATTTCTGAATCTTTCCNCATTATATTATAAAATNTGGAAANGTCAGTTTATCTC
CCCTCCTCNGTATATCTGATTTGTATANGTANGTTGATGNGCTTCTCTCTACAACATTTCTA
GAAAATAGAAAAAAAAGCACAGAGAAATGTTTAACTGTTTGACTCTTATGATACTTCTTGGA
AACTATGACATCAAAGATAGACTTTTGCCTAAGTGGCTTAGCTGGGTCTTTCATAGCCAAAC
TTGTATATTTAATTCTTTGTAATAATAA

FIGURE 45

MPLPWSLALPLLLSWVAGGFGNAASARHHGLLASARQPGVCHYGTKLACCYGWRRNSKGVCE
ATCEPGCKFGECVGPNKCRCFPGYTGKTCSQDVNECGMKPRPCQHRCVNTHGSYKCFCLSGH
MLMPDATCVNSRTCAMINCQYSCEDTEEGPQCLCPSSGLRLAPNGRDCLDIDECASGKVICP
YNRRCVNTFGSYYCKCHIGFELQYISGRYDCIDINECTMDSHTCSHHANCFNTQGSFKCKCK
QGYKGNGLRCSAIPENSVKEVLRAPGTIKDRIKKLLAHKNSMKKKAKIKNVTPEPTRTPTPK
VNLQPFNYEEIVSRGGNSHGGKKGNEEK

Signal peptide:
amino acids 1-21

EGF-like domain cysteine pattern signature.
amino acids 80-91

Calcium-binding EGF-like domains
amino acids 103-124, 230-251 and 185-206

FIGURE 46

GGGAGCTGCTGCTGTGGCTGCTGGTGCTGTGCGCGCTGCTCCTGCTCTTGGTGCAGCTGCTG
CGCTTCCTGAGGGCTGACGGCGACCTGACGCTACTATGGGCCGAGTGGCAGGGACGACGCCC
AGAATGGGAGCTGACTGATATGGTGGTGTGGGTGACTGGAGCCTCGAGTGGAATTGGTGAGG
AGCTGGCTTACCAGTTGTCTAAACTAGGAGTTTCTCTTGTGCTGTCAGCCAGAAGAGTGCAT
GAGCTGGAAAGGGTGAAAAGAAGATGCCTAGAGAATGGCAATTTAAAAGAAAAAGATATACT
TGTTTTGCCCCTTGACCTGACCGACACTGGTTCCCATGAAGCGGCTACCAAAGCTGTTCTCC
AGGAGTTTGGTAGAATCGACATTCTGGTCAACAATGGTGGAATGTCCCAGCGTTCTCTGTGC
ATGGATACCAGCTTGGATGTCTACAGAAAGCTAATAGAGCTTAACTACTTAGGGACGGTGTC
CTTGACAAAATGTGTTCTGCCTCACATGATCGAGAGGAAGCAAGGAAAGATTGTTACTGTGA
ATAGCATCCTGGGTATCATATCTGTACCTCTTTCCATTGGATACTGTGCTAGCAAGCATGCT
CTCCGGGGTTTTTTTAATGGCCTTCGAACAGAACTTGCCACATACCCAGGTATAATAGTTTC
TAACATTTGCCCAGGACCTGTGCAATCAAATATTGTGGAGAATTCCCTAGCTGGAGAAGTCA
CAAAGACTATAGGCAATAATGGAGACCAGTCCCACAAGATGACAACCAGTCGTTGTGTGCGG
CTGATGTTAATCAGCATGGCCAATGATTTGAAAGAAGTTTGGATCTCAGAACAACCTTTCTT
GTTAGTAACATATTTGTGGCAATACATGCCAACCTGGGCCTGGTGGATAACCAACAAGATGG
GGAAGAAAAGGATTGAGAACTTTAAGAGTGGTGTGGATGCAGACTCTTCTTATTTTAAAATC
TTTAAGACAAAACATGACTGAAAAGAGCACCTGTACTTTTCAAGCCACTGGAGGGAGAAATG
GAAAACATGAAAACAGCAATCTTCTTATGCTTCTGAATAATCAAAGACTAATTTGTGATTTT
ACTTTTTAATAGATATGACTTTGCTTCCAACATGGAATGAAATAAAAAATAAATAATAAAAG
ATTGCCATGAATCTTGCAAAA

FIGURE 47

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA36343
><subunit 1 of 1, 289 aa, 1 stop
><MW: 32268, pI: 9.21, NX(S/T): 0
MVVWVTGASSGIGEELAYQLSKLGVSLVLSARRVHELERVKRRCLENGNLKEKDILVLPLDL
TDTGSHEAATKAVLQEFGRIDILVNNGGMSQRSLCMDTSLDVYRKLIELNYLGTVSLTKCVL
PHMIERKQGKIVTVNSILGIISVPLSIGYCASKHALRGFFNGLRTELATYPGIIVSNICPGP
VQSNIVENSLAGEVTKTIGNNGDQSHKMTTSRCVRLMLISMANDLKEVWISEQPFLLVTYLW
QYMPTWAWWITNKMGKKRIENFKSGVDADSSYFKIFKTKHD

Important Features:

Signal Peptide:
amino acids 1-31

Transmembrane domain:
amino acids 136-157

Tyrosine kinase phosphorylation site.
106-113 and 107-114

Homologous region to Short-chain alcohol dehydrogenase
amino acids 80-90, 131-168, 1-13 and 176-185

FIGURE 48

GCGACGTGGGCACCGCCATCAGCTGTTCGCGCGTCTTCTCCTCCAGGTGGGGCAGGGGTTTC
GGGCTGGTGGAGCATGTGCTGGGACAGGACAGCATCCTCAATCAATCCAACAGCATATTCGG
TTGCATCTTCTACACACTACAGCTATTGTTAGGTTGCCTGCGGACACGCTGGGCCTCTGTCC
TGATGCTGCTGAGCTCCCTGGTGTCTCTCGCTGGTTCTGTCTACCTGGCCTGGATCCTGTTC
TTCGTGCTCTATGATTTCTGCATTGTTTGTATCACCACCTATGCTATCAACGTGAGCCTGAT
GTGGCTCAGTTTCCGGAAGGTCCAAGAACCCCAGGGCAAGGCTAAGAGGCACTGAGCCCTCA
ACCCAAGCCAGGCTGACCTCATCTGCTTTGCTTTGGTCTTCAAGCCGCTCAGCGTGCCTGTG
GACAGCGTGGCCCCGGCCCCCCCAAGCCTCAGGAGGGCAACACAGTCCCTGGCGAGTGGCCC
TGGCAGGCCAGTGTGAGGAGGCAAGGAGCCCACATCTGCAGCGGCTCCCTGGTGGCAGACAC
CTGGGTCCTCACTGCTGCCCACTGCTTTGAAAAGGCAGCAGCAACAGAACTGAATTCCTGGT
CAGTGGTCCTGGGTTCTCTGCAGCGTGAGGGACTCAGCCCTGGGGCCGAAGAGGTGGGGGTG
GCTGCCCTGCAGTTGCCCAGGGCCTATAACCACTACAGCCAGGGCTCAGACCTGGCCCTGCT
GCAGCTCGCCCACCCCACGACCCACACACCCCTCTGCCTGCCCCAGCCCGCCCATCGCTTCC
CCTTTGGAGCCTCCTGCTGGGCCACTGGCTGGGATCAGGACACCAGTGATGCTCCTGGGACC
CTACGCAATCTGCGCCTGCGTCTCATCAGTCGCCCCACATGTAACTGTATCTACAACCAGCT
GCACCAGCGACACCTGTCCAACCCGGCCCGGCCTGGGATGCTATGTGGGGGCCCCCAGCCTG
GGGTGCAGGGCCCCTGTCAGGGAGATTCCGGGGGCCCTGTGCTGTGCCTCGAGCCTGACGGA
CACTGGGTTCAGGCTGGCATCATCAGCTTTGCATCAAGCTGTGCCCAGGAGGACGCTCCTGT
GCTGCTGACCAACACAGCTGCTCACAGTTCCTGGCTGCAGGCTCGAGTTCAGGGGGCAGCTT
TCCTGGCCCAGAGCCCAGAGACCCCGGAGATGAGTGATGAGGACAGCTGTGTAGCCTGTGGA
TCCTTGAGGACAGCAGGTCCCCAGGCAGGAGCACCCTCCCCATGGCCCTGGGAGGCCAGGCT
GATGCACCAGGGACAGCTGGCCTGTGGCGGAGCCCTGGTGTCAGAGGAGGCGGTGCTAACTG
CTGCCCACTGCTTCATTGGGCGCCAGGCCCCAGAGGAATGGAGCGTAGGGCTGGGGACCAGA
CCGGAGGAGTGGGGCCTGAAGCAGCTCATCCTGCATGGAGCCTACACCCACCCTGAGGGGGG
CTACGACATGGCCCTCCTGCTGCTGGCCCAGCCTGTGACACTGGGAGCCAGCCTGCGGCCCC
TCTGCCTGCCCTATCCTGACCACCACCTGCCTGATGGGGAGCGTGGCTGGGTTCTGGGACGG
GCCCGCCCAGGAGCAGGCATCAGCTCCCTCCAGACAGTGCCCGTGACCCTCCTGGGGCCTAG
GGCCTGCAGCCGGCTGCATGCAGCTCCTGGGGGTGATGGCAGCCCTATTCTGCCGGGGATGG
TGTGTACCAGTGCTGTGGGTGAGCTGCCCAGCTGTGAGGGCCTGTCTGGGGCACCACTGGTG
CATGAGGTGAGGGGCACATGGTTCCTGGCCGGGCTGCACAGCTTCGGAGATGCTTGCCAAGG
CCCCGCCAGGCCGGCGGTCTTCACCGCGCTCCCTGCCTATGAGGACTGGGTCAGCAGTTTGG
ACTGGCAGGTCTACTTCGCCGAGGAACCAGAGCCCGAGGCTGAGCCTGGAAGCTGCCTGGCC
AACATAAGCCAACCAACCAGCTGCTGACAGGGGACCTGGCCATTCTCAGGACAAGAGAATGC
AGGCAGGCAAATGGCATTACTGCCCCTGTCCTCCCCACCCTGTCATGTGTGATTCCAGGCAC
CAGGGCAGGCCCAGAAGCCCAGCAGCTGTGGGAAGGAACCTGCCTGGGGCCACAGGTGCCCA
CTCCCCACCCTGCAGGACAGGGGTGTCTGTGGACACTCCCACACCCAACTCTGCTACCAAGC
AGGCGTCTCAGCTTTCCTCCTCCTTTACTCTTTCAGATACAATCACGCCAGCCACGTTGTTT
TGAAAATTTCTTTTTTTGGGGGGCAGCAGTTTTCCTTTTTTTAAACTTAAATAAATTGTTAC
AAAATAAAA

FIGURE 49

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA40571
MLLSSLVSLAGSVYLAWILFFVLYDFCIVCITTYAINVSLMWLSFRKVQEPQGKAKRHGNTV
PGEWPWQASVRRQGAHICSGSLVADTWVLTAAHCFEKAAATELNSWSVVLGSLQREGLSPGA
EEVGVAALQLPRAYNHYSQGSDLALLQLAHPTTHTPLCLPQPAHRFPFGASCWATGWDQDTS
DAPGTLRNLRLRLISRPTCNCIYNQLHQRHLSNPARPGMLCGGPQPGVQGPCQGDSGGPVLC
LEPDGHWVQAGIISFASSCAQEDAPVLLTNTAAHSSWLQARVQGAAFLAQSPETPEMSDEDS
CVACGSLRTAGPQAGAPSPWPWEARLMHQGQLACGGALVSEEAVLTAAHCFIGRQAPEEWSV
GLGTRPEEWGLKQLILHGAYTHPEGGYDMALLLLAQPVTLGASLRPLCLPYPDHHLPDGERG
WVLGRARPGAGISSLQTVPVTLLGPRACSRLHAAPGGDGSPILPGMVCTSAVGELPSCEGLS
GAPLVHEVRGTWFLAGLHSFGDACQGPARPAVFTALPAYEDWVSSLDWQVYFAEEPEPEAEP
GSCLANISQPTSC

Important features:

Signal peptide:
amino acids 1-15

Homologous region to Serine proteases, trypsin family
amino acids 79-95, 343-359 and 237-247

N-glycosylation sites.
amino acids 37-40 and 564-567

Kringle domains
amino acids 79-96, 343-360 and 235-247

FIGURE 50

CGGGCCGCCCCCGGCCCCCATTCGGGCCGGGCCTCGCTGCGGCGGCGACTGAGCCAGGCTGG
GCCGCGTCCCTGAGTCCCAGAGTCGGCGCGGCGCGGCAGGGGCAGCCTTCCACCACGGGGAG
CCCAGCTGTCAGCCGCCTCACAGGAAGATGCTGCGTCGGCGGGGCAGCCCTGGCATGGGTGT
GCATGTGGGTGCAGCCCTGGGAGCACTGTGGTTCTGCCTCACAGGAGCCCTGGAGGTCCAGG
TCCCTGAAGACCCAGTGGTGGCACTGGTGGGCACCGATGCCACCCTGTGCTGCTCCTTCTCC
CCTGAGCCTGGCTTCAGCCTGGCACAGCTCAACCTCATCTGGCAGCTGACAGATACCAAACA
GCTGGTGCACAGCTTTGCTGAGGGCCAGGACCAGGGCAGCGCCTATGCCAACCGCACGGCCC
TCTTCCCGGACCTGCTGGCACAGGGCAACGCATCCCTGAGGCTGCAGCGCGTGCGTGTGGCG
GACGAGGGCAGCTTCACCTGCTTCGTGAGCATCCGGGATTTCGGCAGCGCTGCCGTCAGCCT
GCAGGTGGCCGCTCCCTACTCGAAGCCCAGCATGACCCTGGAGCCCAACAAGGACCTGCGGC
CAGGGGACACGGTGACCATCACGTGCTCCAGCTACCAGGGCTACCCTGAGGCTGAGGTGTTC
TGGCAGGATGGGCAGGGTGTGCCCCTGACTGGCAACGTGACCACGTCGCAGATGGCCAACGA
GCAGGGCTTGTTTGATGTGCACAGCGTCCTGCGGGTGGTGCTGGGTGCGAATGGCACCTACA
GCTGCCTGGTGCGCAACCCCGTGCTGCAGCAGGATGCGCACRGCTCTGTCACCATCACAGGG
CAGCCTATGACATTCCCCCCAGAGGCCCTGTGGGTGACCGTGGGGCTGTCTGTCTGTCTCAT
TGCACTGCTGGTGGCCCTGGCTTTCGTGTGCTGGAGAAAGATCAAACAGAGCTGTGAGGAGG
AGAATGCAGGAGCTGAGGACCAGGATGGGGAGGGAGAAGGCTCCAAGACAGCCCTGCAGCCT
CTGAAACACTCTGACAGCAAAGAAGATGATGGACAAGAAATAGCCTGACCATGAGGACCAGG
GAGCTGCTACCCCTCCCTACAGCTCCTACCCTCTGGCTGCAATGGGGCTGCACTGTGAGCCC
TGCCCCCAACAGATGCATCCTGCTCTGACAGGTGGGCTCCTTCTCCAAAGGATGCGATACAC
AGACCACTGTGCAGCCTTATTTCTCCAATGGACATGATTCCCAAGTCATCCTGCTGCCTTTT
TTCTTATAGACACAATGAACAGACCACCCACAACCTTAGTTCTCTAAGTCATCCTGCCTGCT
GCCTTATTTCACAGTACATACATTTCTTAGGGACACAGTACACTGACCACATCACCACCCTC
TTCTTCCAGTGCTGCGTGGACCATCTGGCTGCCTTTTTTCTCCAAAAGATGCAATATTCAGA
CTGACTGACCCCCTGCCTTATTTCACCAAAGACACGATGCATAGTCACCCCGGCCTTGTTTC
TCCAATGGCCGTGATACACTAGTGATCATGTTCAGCCCTGCTTCCACCTGCATAGAATCTTT
TCTTCTCAGACAGGGACAGTGCGGCCTCAACATCTCCTGGAGTCTAGAAGCTGTTTCCTTTC
CCCTCCTTCCTCCCTGCCCCAAGTGAAGACAGGGCAGGGCCAGGAATGCTTTGGGGACACCG
AGGGGACTGCCCCCCACCCCCACCATGGTGCTATTCTGGGGCTGGGGCAGTCTTTTCCTGGC
TTGCCTCTGGCCAGCTCCTGGCCTCTGGTAGAGTGAGACTTCAGACGTTCTGATGCCTTCCG
GATGTCATCTCTCCCTGCCCCAGGAATGGAAGATGTGAGGACTTCTAATTTAAATGTGGGAC
TCGGAGGGATTTTGTAAACTGGGGGTATATTTTGGGGAAAATAAATGTCTTTGTAAAAAAAA
AAAAAAAAAAAAAA

FIGURE 51

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA41386
><subunit 1 of 1, 316 aa, 1 stop, 1 unknown
><MW: -1, pI: 4.62, NX(S/T): 4
MLRRRGSPGMGVHVGAALGALWFCLTGALEVQVPEDPVVALVGTDATLCCSFSPEPGFSLAQ
LNLIWQLTDTKQLVHSFAEGQDQGSAYANRTALFPDLLAQGNASLRLQRVRVADEGSFTCFV
SIRDFGSAAVSLQVAAPYSKPSMTLEPNKDLRPGDTVTITCSSYQGYPEAEVFWQDGQGVPL
TGNVTTSQMANEQGLFDVHSVLRVVLGANGTYSCLVRNPVLQQDAHXSVTITGQPMTFPPEA
LWVTVGLSVCLIALLVALAFVCWRKIKQSCEEENAGAEDQDGEGEGSKTALQPLKHSDSKED
DGQEIA

Important features:
Signal peptide:
amino acids 1-28

Transmembrane domain:
amino acids 251-270

N-glycosylation site.
amino acids 91-94, 104-107, 189-192 and 215-218

Homologous region to Immunoglobulins and MHC
amino acids 217-234

FIGURE 52

TTCGTGACCCTTGAGAAAAGAGTTGGTGGTAAATGTGCCACGTCTTCTAAGAAGGGGGAGTC
CTGAACTTGTCTGAAGCCCTTGTCCGTAAGCCTTGAACTACGTTCTTAAATCTATGAAGTCG
AGGGACCTTTCGCTGCTTTTGTAGGGACTTCTTTCCTTGCTTCAGCAACATGAGGCTTTTCT
TGTGGAACGCGGTCTTGACTCTGTTCGTCACTTCTTTGATTGGGGCTTTGATCCCTGAACCA
GAAGTGAAAATTGAAGTTCTCCAGAAGCCATTCATCTGCCATCGCAAGACCAAAGGAGGGGA
TTTGATGTTGGTCCACTATGAAGGCTACTTAGAAAAGGACGGCTCCTTATTTCACTCCACTC
ACAAACATAACAATGGTCAGCCCATTTGGTTTACCCTGGGCATCCTGGAGGCTCTCAAAGGT
TGGGACCAGGGCTTGAAAGGAATGTGTGTAGGAGAGAAGAGAAAGCTCATCATTCCTCCTGC
TCTGGGCTATGGAAAAGAAGGAAAAGGTAAAATTCCCCCAGAAAGTACACTGATATTTAATA
TTGATCTCCTGGAGATTCGAAATGGACCAAGATCCCATGAATCATTCCAAGAAATGGATCTT
AATGATGACTGGAAACTCTCTAAAGATGAGGTTAAAGCATATTTAAAGAAGGAGTTTGAAAA
ACATGGTGCGGTGGTGAATGAAAGTCATCATGATGCTTTGGTGGAGGATATTTTTGATAAAG
AAGATGAAGACAAAGATGGGTTTATATCTGCCAGAGAATTTACATATAAACACGATGAGTTA
TAGAGATACATCTACCCTTTTAATATAGCACTCATCTTTCAAGAGAGGGCAGTCATCTTTAA
AGAACATTTTATTTTTATACAATGTTCTTTCTTGCTTTGTTTTTTATTTTTATATATTTTTT
CTGACTCCTATTTAAAGAACCCCTTAGGTTTCTAAGTACCCATTTCTTTCTGATAAGTTATT
GGGAAGAAAAAGCTAATTGGTCTTTGAATAGAAGACTTCTGGACAATTTTTCACTTTCACAG
ATATGAAGCTTTGTTTTACTTTCTCACTTATAAATTTAAAATGTTGCAACTGGGAATATACC
ACGACATGAGACCAGGTTATAGCACAAATTAGCACCCTATATTTCTGCTTCCCTCTATTTTC
TCCAAGTTAGAGGTCAACATTTGAAAAGCCTTTTGCAATAGCCCAAGGCTTGCTATTTTCAT
GTTATAATGAAATAGTTTATGTGTAACTGGCTCTGAGTCTCTGCTTGAGGACCAGAGGAAAA
TGGTTGTTGGACCTGACTTGTTAATGGCTACTGCTTTACTAAGGAGATGTGCAATGCTGAAG
TTAGAAACAAGGTTAATAGCCAGGCATGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGAG
GCTGAGGCGGGCGGATCACCTGAGGTTGGGAGTTCGAGACCAGCCTGACCAACACGGAGAAA
CCCTATCTCTACTAAAAATACAAAGTAGCCCGGCGTGGTGATGCGTGCCTGTAATCCCAGCT
ACCCAGGAAGGCTGAGGCGGCAGAATCACTTGAACCCGAGGCCGAGGTTGCGGTAAGCCGAG
ATCACCTNCAGCCTGGACACTCTGTCTCGAAAAAAGAAAAGAACACGGTTAATACCATATNA
ATATGTATGCATTGAGACATGCTACCTAGGACTTAAGCTGATGAAGCTTGGCTCCTAGTGAT
TGGTGGCCTATTATGATAAATAGGACAAATCATTTATGTGTGAGTTTCTTTGTAATAAAATG
TATCAATATGTTATAGATGAGGTAGAAAGTTATATTTATATTCAATATTTACTTCTTAAGGC
TAGCGGAATATCCTTCCTGGTTCTTTAATGGGTAGTCTATAGTATATTATACTACAATAACA
TTGTATCATAAGATAAAGTAGTAAACCAGTCTACATTTTCCCATTTCTGTCTCATCAAAAAC
TGAAGTTAGCTGGGTGTGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGGGCCAAGGAGGG
TGGATCACTTGAGATCAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACCTTGTCTCTA
CTAAAAATACAAAAATTAGCCAGGCGTGGTGGTGCACACCTGTAGTCCCAGCTACTCGGGAG
GCTGAGACAGGAGATTTGCTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCCAAGATTGTGCC
ACTGCACTCCAGCCTGGGTGACAGAGCAAGACTCCATCTCAAAAAAAAAAAAAAGAAGCAGA
CCTACAGCAGCTACTATTGAATAAATACCTATCCTGGATTTT

FIGURE 53

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44194
><subunit 1 of 1, 211 aa, 1 stop
><MW: 24172, pI: 5.99, NX(S/T): 1
MRLFLWNAVLTLFVTSLIGALIPEPEVKIEVLQKPFICHRKTKGGDLMLVHYEGYLEKDGSL
FHSTHKHNNGQPIWFTLGILEALKGWDQGLKGMCVGEKRKLIIPPALGYGKEGKGKIPPEST
LIFNIDLLEIRNGPRSHESFQEMDLNDDWKLSKDEVKAYLKKEFEKHGAVVNESHHDALVED
IFDKEDEDKDGFISAREFTYKHDEL

Important features:
Signal peptide:
amino acids 1-20

N-glycosylation site.
amino acids 176-179

Casein kinase II phosphorylation site.
amino acids 143-146, 156-159, 178-181 and 200-203

Endoplasmic reticulum targeting sequence.
amino acids 208-211

FKBP-type peptidyl-prolyl cis-trans isomerase
amino acids 78-114 and 118-131

EF-hand calcium-binding domain.
amino acids 191-203, 184-203 and 140-159

S-100/ICaBP type calcium binding domain
amino acids 183-203

FIGURE 54

AATAAAGCTTCCTTAATGTTGTATATGTCTTTGAAGTACATCCGTGCATTTTTTTTTAGCAT
CCAACCATTCCTCCCTTGTAGTTCTCGCCCCCTCAAATCACCCTCTCCCGTAGCCCACCCGA
CTAACATCTCAGTCTCTGAAAATGCACAGAGATGCCTGGCTACCTCGCCCTGCCTTCAGCCT
CACGGGGCTCAGTCTCTTTTTCTCTTTGGTGCCACCAGGACGGAGCATGGAGGTCACAGTAC
CTGCCACCCTCAACGTCCTCAATGGCTCTGACGCCCGCCTGCCCTGCACCTTCAACTCCTGC
TACACAGTGAACCACAAACAGTTCTCCCTGAACTGGACTTACCAGGAGTGCAACAACTGCTC
TGAGGAGATGTTCCTCCAGTTCCGCATGAAGATCATTAACCTGAAGCTGGAGCGGTTTCAAG
ACCGCGTGGAGTTCTCAGGGAACCCCAGCAAGTACGATGTGTCGGTGATGCTGAGAAACGTG
CAGCCGGAGGATGAGGGGATTTACAACTGCTACATCATGAACCCCCCTGACCGCCACCGTGG
CCATGGCAAGATCCATCTGCAGGTCCTCATGGAAGAGCCCCCTGAGCGGGACTCCACGGTGG
CCGTGATTGTGGGTGCCTCCGTCGGGGGCTTCCTGGCTGTGGTCATCTTGGTGCTGATGGTG
GTCAAGTGTGTGAGGAGAAAAAAAGAGCAGAAGCTGAGCACAGATGACCTGAAGACCGAGGA
GGAGGGCAAGACGGACGGTGAAGGCAACCCGGATGATGGCGCCAAGTAGTGGGTGGCCGGCC
CTGCAGCCTCCCGTGTCCCGTCTCCTCCCCTCTCCGCCCTGTACAGTGACCCTGCCTGCTCG
CTCTTGGTGTGCTTCCCGTGACCTAGGACCCCAGGGCCCACCTGGGGCCTCCTGAACCCCCG
ACTTCGTATCTCCCACCCTGCACCAAGAGTGACCCACTCTCTTCCATCCGAGAAACCTGCCA
TGCTCTGGGACGTGTGGGCCCTGGGGAGAGGAGAGAAAGGGCTCCCACCTGCCAGTCCCTGG
GGGGAGGCAGGAGGCACATGTGAGGGTCCCCAGAGAGAAGGGAGTGGGTGGGCAGGGGTAGA
GGAGGGGCCGCTGTCACCTGCCCAGTGCTTGCCTGGCAGTGGCTTCAGAGAGGACCTGGTGG
GGAGGGAGGGCTTTCCTGTGCTGACAGCGCTCCCTCAGGAGGGCCTTGGCCTGGCACGGCTG
TGCTCCTCCCCTGCTCCCAGCCCAGAGCAGCCATCAGGCTGGAGGTGACGATGAGTTCCTGA
AACTTGGAGGGGCATGTTAAAGGGATGACTGTGCATTCCAGGGCACTGACGGAAAGCCAGGG
CTGCAGGCAAAGCTGGACATGTGCCCTGGCCCAGGAGGCCATGTTGGGCCCTCGTTTCCATT
GCTAGTGGCCTCCTTGGGGCTCCTGTTGGCTCCTAATCCCTTAGGACTGTGGATGAGGCCAG
ACTGGAAGAGCAGCTCCAGGTAGGGGGCCATGTTTCCCAGCGGGGACCCACCAACAGAGGCC
AGTTTCAAAGTCAGCTGAGGGGCTGAGGGGTGGGGCTCCATGGTGAATGCAGGTTGCTGCAG
GCTCTGCCTTCTCCATGGGGTAACCACCCTCGCCTGGGCAGGGGCAGCCAAGGCTGGGAAAT
GAGGAGGCCATGCACAGGGTGGGGCAGCTTTCTTTGGGGCTTCAGTGAGAACTCTCCCAGTT
GCCCTTGGTGGGGTTTCCACCTGGCTTTTGGCTACAGAGAGGGAAGGGAAAGCCTGAGGCCG
GCATAAGGGGAGGCCTTGGAACCTGAGCTGCCAATGCCAGCCCTGTCCCATCTGCGGCCACG
CTACTCGCTCCTCTCCCAACAACTCCCTTCGTGGGGACAAAAGTGACAATTGTAGGCCAGGC
ACAGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGCGGGTGGATTACCTCCAT
CTGTTTAGTAGAAATGGGCAAAACCCCATCTCTACTAAAAATACAAGAATTAGCTGGGCGTG
GTGGCGTGTGCCTGTAATCCCAGCTATTTGGGAGGCTGAGGCAGGAGAATCGCTTGAGCCCG
GGAAGCAGAGGTTGCAGTGAACTGAGATAGTGATAGTGCCACTGCAATTCAGCCTGGGTGAC
ATAGAGAGACTCCATCTCAAAAAAAA

FIGURE 55

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45415
<subunit 1 of 1, 215 aa, 1 stop
<MW: 24326, pI: 6.32, NX(S/T): 4
MHRDAWLPRPAFSLTGLSLFFSLVPPGRSMEVTVPATLNVLNGSDARLPCTFNSCYTVNHKQ
FSLNWTYQECNNCSEEMFLQFRMKIINLKLERFQDRVEFSGNPSKYDVSVMLRNVQPEDEGI
YNCYIMNPPDRHRGHGKIHLQVLMEEPPERDSTVAVIVGASVGGFLAVVILVLMVVKCVRRK
KEQKLSTDDLKTEEEGKTDGEGNPDDGAK

Important features:

Signal peptide:
amino acids 1-20

Transmembrane domain:
amino acids 161-179

Immunoglobulin-like fold:
amino acids 83-127

N-glycosylation sites.
amino acids 42-45, 66-69 and 74-77

FIGURE 56

GTTGTATATGTCCTGAAGTACATCCGTGCATTTTTTTTAGCATCCAACCATCCTCCCTTGTA
GTTCTCGCCCCCTCAAATCACCTTCTCCCTTAGCCCACCCNACTAACATCTCAGTCTCTGAA
AATGCACAGAGATGCCTGGCTACCTCGCCCTGCCTTCAGCCTCACGGGGCTCAGTCTCTTTT
TCTCTTTGGTGCCACCAGGACGGAGCATGGAGGTCCACAGTACCTGNCCACCCTCAACGTCC
TCAATGGCTCTGACGCCCGCCTGCCCTGCCCTTCAACTCCTGCTACACAGTGAACCACAAAC
AGTTCTCCCTGAACTGGACTTACCAGGAGTGCAACAACTGCTCTGAGGAGATGTTCCTCCAG
TTCCGCATGAAGATCATTAACCTGAAGCTGGAGCGGTTTCAAGACCGCGTGGAGTTCTCAGG
GAACCCCAGCAAGTACGATGTGTCGGTGATGCTGAGAAACGTGCAGCCGGAGGATGAGGGGA
TTTACAACTGCTACATCATGAACCCCCC

FIGURE 57

TCACGGGGCTCATCTCTTTTTCTCTTTGGTGCCCACCAGGACGGAGCATGGAGGTNCACATA
CCTGCCACCCTCAACGTCCTCAATGGCTTTGACGCCCGCCTGCCCTGCACCTTCAACTCCNG
CTACACAGTGAACCACAAACAGTTCTCCCTGAACTGGATTTACCAGGAGTGCAACAACTGGC
TCTGAGGAGATGTTCCTCCAGTTCCCGCATGGAAGATCATTTAACCTGAAAGCTGGAAGCGG
TTTTCAAGAACCGCGTGGAAGTTTCTCAGGGAACCCCAGCAAGTACGATGTGTCGGTGATGC
TGAGAAACGTGCAGCCGGAGGATGAGGGGATTTACAACTGCTACATCATGAACCCCCC

FIGURE 58

TGCGGCGACCGTCGTACACC**ATG**GGCCTCCACCTCCGCCCCTACCGTGTGGGGCTGCTCCCG
GATGGCCTCCTGTTCCTCTTGCTGCTGCTAATGCTGCTCGCGGACCCAGCGCTCCCGGCCGG
ACGTCACCCCCCAGTGGTGCTGGTCCCTGGTGATTTGGGTAACCAACTGGAAGCCAAGCTGG
ACAAGCCGACAGTGGTGCACTACCTCTGCTCCAAGAAGACCGAAAGCTACTTCACAATCTGG
CTGAACCTGGAACTGCTGCTGCCTGTCATCATTGACTGCTGGATTGACAATATCAGGCTGGT
TTACAACAAAACATCCAGGGCCACCCAGTTTCCTGATGGTGTGGATGTACGTGTCCCTGGCT
TTGGGAAGACCTTCTCACTGGAGTTCCTGGACCCCAGCAAAAGCAGCGTGGGTTCCTATTTC
CACACCATGGTGGAGAGCCTTGTGGGCTGGGGCTACACACGGGGTGAGGATGTCCGAGGGGC
TCCCTATGACTGGCGCCGAGCCCCAAATGAAAACGGGCCCTACTTCCTGGCCCTCCGCGAGA
TGATCGAGGAGATGTACCAGCTGTATGGGGGCCCCGTGGTGCTGGTTGCCCACAGTATGGGC
AACATGTACACGCTCTACTTTCTGCAGCGGCAGCCGCAGGCCTGGAAGGACAAGTATATCCG
GGCCTTCGTGTCACTGGGTGCGCCCTGGGGGGGCGTGGCCAAGACCCTGCGCGTCCTGGCTT
CAGGAGACAACAACCGGATCCCAGTCATCGGGCCCCTGAAGATCCGGGAGCAGCAGCGGTCA
GCTGTCTCCACCAGCTGGCTGCTGCCCTACAACTACACATGGTCACCTGAGAAGGTGTTCGT
GCAGACACCCACAATCAACTACACACTGCGGGACTACCGCAAGTTCTTCCAGGACATCGGCT
TTGAAGATGGCTGGCTCATGCGGCAGGACACAGAAGGGCTGGTGGAAGCCACGATGCCACCT
GGCGTGCAGCTGCACTGCCTCTATGGTACTGGCGTCCCCACACCAGACTCCTTCTACTATGA
GAGCTTCCCTGACCGTGACCCTAAAATCTGCTTTGGTGACGGCGATGGTACTGTGAACTTGA
AGAGTGCCCTGCAGTGCCAGGCCTGGCAGAGCCGCCAGGAGCACCAAGTGTTGCTGCAGGAG
CTGCCAGGCAGCGAGCACATCGAGATGCTGGCCAACGCCACCACCCTGGCCTATCTGAAACG
TGTGCTCCTTGGGCCC**TGA**CTCCTGTGCCACAGGACTCCTGTGGCTCGGCCGTGGACCTGCT
GTTGGCCTCTGGGGCTGTCATGGCCCACGCGTTTTGCAAAGTTTGTGACTCACCATTCAAGG
CCCCGAGTCTTGGACTGTGAAGCATCTGCCATGGGGAAGTGCTGTTTGTTATCCTTTCTCTG
TGGCAGTGAAGAAGGAAGAAATGAGAGTCTAGACTCAAGGGACACTGGATGGCAAGAATGCT
GCTGATGGTGGAACTGCTGTGACCTTAGGACTGGCTCCACAGGGTGGACTGGCTGGGCCCTG
GTCCCAGTCCCTGCCTGGGGCCATGTGTCCCCCTATTCCTGTGGGCTTTTCATACTTGCCTA
CTGGGCCCTGGCCCCGCAGCCTTCCTATGAGGGATGTTACTGGGCTGTGGTCCTGTACCCAG
AGGTCCCAGGGATCGGCTCCTGGCCCCTCGGGTGACCCTTCCCACACACCAGCCACAGATAG
GCCTGCCACTGGTCATGGGTAGCTAGAGCTGCTGGCTTCCCTGTGGCTTAGCTGGTGGCCAG
CCTGACTGGCTTCCTGGGCGAGCCTAGTAGCTCCTGCAGGCAGGGGCAGTTTGTTGCGTTCT
TCGTGGTTCCCAGGCCCTGGGACATCTCACTCCACTCCTACCTCCCTTACCACCAGGAGCAT
TCAAGCTCTGGATTGGGCAGCAGATGTGCCCCCAGTCCCGCAGGCTGTGTTCCAGGGGCCCT
GATTTCCTCGGATGTGCTATTGGCCCCAGGACTGAAGCTGCCTCCCTTCACCCTGGGACTGT
GGTTCCAAGGATGAGAGCAGGGGTTGGAGCCATGGCCTTCTGGGAACCTATGGAGAAAGGGA
ATCCAAGGAAGCAGCCAAGGCTGCTCGCAGCTTCCCTGAGCTGCACCTCTTGCTAACCCCAC
CATCACACTGCCACCCTGCCCTAGGGTCTCACTAGTACCAAGTGGGTCAGCACAGGGCTGAG
GATGGGGCTCCTATCCACCCTGGCCAGCACCCAGCTTAGTGCTGGGACTAGCCCAGAAACTT
GAATGGGACCCTGAGAGAGCCAGGGGTCCCCTGAGGCCCCCCTAGGGGCTTTCTGTCTGCCC
CAGGGTGCTCCATGGATCTCCCTGTGGCAGCAGGCATGGAGAGTCAGGGCTGCCTTCATGGC
AGTAGGCTCTAAGTGGGTGACTGGCCACAGGCCGAGAAAAGGGTACAGCCTCTAGGTGGGGT
TCCCAAAGACGCCTTCAGGCTGGACTGAGCTGCTCTCCCACAGGGTTTCTGTGCAGCTGGAT
TTTCTCTGTTGCATACATGCCTGGCATCTGTCTCCCCTTGTTCCTGAGTGGCCCCACATGGG
GCTCTGAGCAGGCTGTATCTGGATTCTGGCAATAAAAGTACTCTGGATGCTGTAAAAAAAAA
AAAAAAAAAAAAA

FIGURE 59

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44189
><subunit 1 of 1, 412 aa, 1 stop
><MW: 46658, pI: 6.65, NX(S/T): 4
MGLHLRPYRVGLLPDGLLFLLLLLMLLADPALPAGRHPPVVLVPGDLGNQLEAKLDKPTVVH
YLCSKKTESYFTIWLNLELLLPVIIDCWIDNIRLVYNKTSRATQFPDGVDVRVPGFGKTFSL
EFLDPSKSSVGSYFHTMVESLVGWGYTRGEDVRGAPYDWRRAPNENGPYFLALREMIEEMYQ
LYGGPVVLVAHSMGNMYTLYFLQRQPQAWKDKYIRAFVSLGAPWGGVAKTLRVLASGDNNRI
PVIGPLKIREQQRSAVSTSWLLPYNYTWSPEKVFVQTPTINYTLRDYRKFFQDIGFEDGWLM
RQDTEGLVEATMPPGVQLHCLYGTGVPTPDSFYYESFPDRDPKICFGDGDGTVNLKSALQCQ
AWQSRQEHQVLLQELPGSEHIEMLANATTLAYLKRVLLGP

Important features:

Signal peptide:
amino acids 1-28

Potential lipid substrate binding site:
amino acids 147-164

N-glycosylation sites.
amino acids 99-102, 273-276, 289-292 and 398-401

Lipases, serine proteins
amino acids 189-201

Beta-transducin family Trp-Asp repeat
amino acids 353-365

FIGURE 60

CGGACGCGTGGGCGGACGCGTGGGGCGGCGGCAGCGGCGGCGACGGCGACATGGAGAGCGGG
GCCTACGGCGCGGCCAAGGCGGGCGGCTCCTTCGACCTGCGGCGCTTCCTGACGCAGCCGCA
GGTGGTGGCGCGCGCCGTGTGCTTGGTCTTCGCCTTGATCGTGTTCTCCTGCATCTATGGTG
AGGGCTACAGCAATGCCCACGAGTCTAAGCAGATGTACTGCGTGTTCAACCGCAACGAGGAT
GCCTGCCGCTATGGCAGTGCCATCGGGGTGCTGGCCTTCCTGGCCTCGGCCTTCTTCTTGGT
GGTCGACGCGTATTTCCCCCAGATCAGCAACGCCACTGACCGCAAGTACCTGGTCATTGGTG
ACCTGCTCTTCTCAGCTCTCTGGACCTTCCTGTGGTTTGTTGGTTTCTGCTTCCTCACCAAC
CAGTGGGCAGTCACCAACCCGAAGGACGTGCTGGTGGGGGCCGACTCTGTGAGGGCAGCCAT
CACCTTCAGCTTCTTTTCCATCTTCTCCTGGGGTGTGCTGGCCTCCCTGGCCTACCAGCGCT
ACAAGGCTGGCGTGGACGACTTCATCCAGAATTACGTTGACCCCACTCCGGACCCCAACACT
GCCTACGCCTCCTACCCAGGTGCATCTGTGGACAACTACCAACAGCCACCCTTCACCCAGAA
CGCGGAGACCACCGAGGGCTACCAGCCGCCCCCTGTGTACTGAGTGGCGGTTAGCGTGGGAA
GGGGGACAGAGAGGGCCCTCCCCTCTGCCCTGGACTTTCCCATCAGCCTCCTGGAACTGCCA
GCCCCTCTCTTTCACCTGTTCCATCCTGTGCAGCTGACACACAGCTAAGGAGCCTCATAGCC
TGGCGGGGGCTGGCAGAGCCACACCCCAAGTGCCTGTGCCCAGAGGGCTTCAGTCAGCCGCT
CACTCCTCCAGGGCACTTTTAGGAAAGGGTTTTTAGCTAGTGTTTTTCCTCGCTTTTAATGA
CCTCAGCCCCGCCTGCAGTGGCTAGAAGCCAGCAGGTGCCCATGTGCTACTGACAAGTGCCT
CAGCTTCCCCCCGGCCCGGGTCAGGCCGTGGGAGCCGCTATTATCTGCGTTCTCTGCCAAAG
ACTCGTGGGGGCCATCACACCTGCCCTGTGCAGCGGAGCCGGACCAGGCTCTTGTGTCCTCA
CTCAGGTTTGCTTCCCCTGTGCCCACTGCTGTATGATCTGGGGGCCACCACCCTGTGCCGGT
GGCCTCTGGGCTGCCTCCCGTGGTGTGAGGGCGGGGCTGGTGCTCATGGCACTTCCTCCTTG
CTCCCACCCCTGGCAGCAGGGAAGGGCTTTGCCTGACAACACCCAGCTTTATGTAAATATTC
TGCAGTTGTTACTTAGGAAGCCTGGGGAGGGCAGGGGTGCCCCATGGCTCCCAGACTCTGTC
TGTGCCGAGTGTATTATAAAATCGTGGGGGAGATGCCCGGCCTGGGATGCTGTTTGGAGACG
GAATAAATGTTTTCTCATTCAAAG

FIGURE 61

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48304
<subunit 1 of 1, 224 aa, 1 stop
<MW: 24810, pI: 4.75, NX(S/T): 1
MESGAYGAAKAGGSFDLRRFLTQPQVVARAVCLVFALIVFSCIYGEGYSNAHESKQMYCVFN
RNEDACRYGSAIGVLAFLASAFFLVVDAYFPQISNATDRKYLVIGDLLFSALWTFLWFVGFC
FLTNQWAVTNPKDVLVGADSVRAAITFSFFSIFSWGVLASLAYQRYKAGVDDFIQNYVDPTP
DPNTAYASYPGASVDNYQQPPFTQNAETTEGYQPPPVY

Important features:

Type II Transmembrane domain:
amino acids 24-43

Other transmembrane domains:
amino acids 74-90, 108-126 and 145-161

N-glycosylation site.
amino acids 97-100

FIGURE 62

GAGCCACCTACCCTGCTCCGAGGCCAGGCCTGCAGGGCCTCATCGGCCAGAGGGTGATCAGT
GAGCAGAAGGATGCCCGTGGCCGAGGCCCCCCAGGTGGCTGGCGGGCAGGGGGACGGAGGTG
ATGGCGAGGAAGCGGAGCCAGAGGGGATGTTCAAGGCCTGTGAGGACTCCAAGAGAAAAGCC
CGGGGCTACCTCCGCCTGGTGCCCCTGTTTGTGCTGCTGGCCCTGCTCGTGCTGGCTTCGGC
GGGGGTGCTACTCTGGTATTTCCTAGGGTACAAGGCGGAGGTGATGGTCAGCCAGGTGTACT
CAGGCAGTCTGCGTGTACTCAATCGCCACTTCTCCCAGGATCTTACCCGCCGGGAATCTAGT
GCCTTCCGCAGTGAAACCGCCAAAGCCCAGAAGATGCTCAAGGAGCTCATCACCAGCACCCG
CCTGGGAACTTACTACAACTCCAGCTCCGTCTATTCCTTTGGGGAGGGACCCCTCACCTGCT
TCTTCTGGTTCATTCTCCAAATCCCCGAGCACCGCCGGCTGATGCTGAGCCCCGAGGTGGTG
CAGGCACTGCTGGTGGAGGAGCTGCTGTCCACAGTCAACAGCTCGGCTGCCGTCCCCTACAG
GGCCGAGTACGAAGTGGACCCCGAGGGCCTAGTGATCCTGGAAGCCAGTGTGAAAGACATAG
CTGCATTGAATTCCACGCTGGGTTGTTACCGCTACAGCTACGTGGGCCAGGGCCAGGTCCTC
CGGCTGAAGGGGCCTGACCACCTGGCCTCCAGCTGCCTGTGGCACCTGCAGGGCCCCAAGGA
CCTCATGCTCAAACTCCGGCTGGAGTGGACGCTGGCAGAGTGCCGGGACCGACTGGCCATGT
ATGACGTGGCCGGGCCCCTGGAGAAGAGGCTCATCACCTCGGTGTACGGCTGCAGCCGCCAG
GAGCCCGTGGTGGAGGTTCTGGCGTCGGGGGCCATCATGGCGGTCGTCTGGAAGAAGGGCCT
GCACAGCTACTACGACCCCTTCGTGCTCTCCGTGCAGCCGGTGGTCTTCCAGGCCTGTGAAG
TGAACCTGACGCTGGACAACAGGCTCGACTCCCAGGGCGTCCTCAGCACCCCGTACTTCCCC
AGCTACTACTCGCCCCAAACCCACTGCTCCTGGCACCTCACGGTGCCCTCTCTGGACTACGG
CTTGGCCCTCTGGTTTGATGCCTATGCACTGAGGAGGCAGAAGTATGATTTGCCGTGCACCC
AGGGCCAGTGGACGATCCAGAACAGGAGGCTGTGTGGCTTGCGCATCCTGCAGCCCTACGCC
GAGAGGATCCCCGTGGTGGCCACGGCCGGGATCACCATCAACTTCACCTCCCAGATCTCCCT
CACCGGGCCCGGTGTGCGGGTGCACTATGGCTTGTACAACCAGTCGGACCCCTGCCCTGGAG
AGTTCCTCTGTTCTGTGAATGGACTCTGTGTCCCTGCCTGTGATGGGGTCAAGGACTGCCCC
AACGGCCTGGATGAGAGAAACTGCGTTTGCAGAGCCACATTCCAGTGCAAAGAGGACAGCAC
ATGCATCTCACTGCCCAAGGTCTGTGATGGGCAGCCTGATTGTCTCAACGGCAGCGATGAAG
AGCAGTGCCAGGAAGGGGTGCCATGTGGGACATTCACCTTCCAGTGTGAGGACCGGAGCTGC
GTGAAGAAGCCCAACCCGCAGTGTGATGGGCGGCCCGACTGCAGGGACGGCTCGGATGAGGA
GCACTGTGACTGTGGCCTCCAGGGCCCCTCCAGCCGCATTGTTGGTGGAGCTGTGTCCTCCG
AGGGTGAGTGGCCATGGCAGGCCAGCCTCCAGGTTCGGGGTCGACACATCTGTGGGGGGGCC
CTCATCGCTGACCGCTGGGTGATAACAGCTGCCCACTGCTTCCAGGAGGACAGCATGGCCTC
CACGGTGCTGTGGACCGTGTTCCTGGGCAAGGTGTGGCAGAACTCGCGCTGGCCTGGAGAGG
TGTCCTTCAAGGTGAGCCGCCTGCTCCTGCACCCGTACCACGAAGAGGACAGCCATGACTAC
GACGTGGCGCTGCTGCAGCTCGACCACCCGGTGGTGCGCTCGGCCGCCGTGCGCCCCGTCTG
CCTGCCCGCGCGCTCCCACTTCTTCGAGCCCGGCCTGCACTGCTGGATTACGGGCTGGGGCG
CCTTGCGCGAGGGCGGCCCCATCAGCAACGCTCTGCAGAAAGTGGATGTGCAGTTGATCCCA
CAGGACCTGTGCAGCGAGGCCTATCGCTACCAGGTGACGCCACGCATGCTGTGTGCCGGCTA
CCGCAAGGGCAAGAAGGATGCCTGTCAGGGTGACTCAGGTGGTCCGCTGGTGTGCAAGGCAC
TCAGTGGCCGCTGGTTCCTGGCGGGGCTGGTCAGCTGGGGCCTGGGCTGTGGCCGGCCTAAC
TACTTCGGCGTCTACACCCGCATCACAGGTGTGATCAGCTGGATCCAGCAAGTGGTGACCTG
AGGAACTGCCCCCCTGCAAAGCAGGGCCCACCTCCTGGACTCAGAGAGCCCAGGGCAACTGC
CAAGCAGGGGGACAAGTATTCTGGCGGGGGGTGGGGGAGAGAGCAGGCCCTGTGGTGGCAGG
AGGTGGCATCTTGTCTCGTCCCTGATGTCTGCTCCAGTGATGGCAGGAGGATGGAGAAGTGC
CAGCAGCTGGGGGTCAAGACGTCCCCTGAGGACCCAGGCCCACACCCAGCCCTTCTGCCTCC
CAATTCTCTCTCCTCCGTCCCCTTCCTCCACTGCTGCCTAATGCAAGGCAGTGGCTCAGCAG
CAAGAATGCTGGTTCTACATCCCGAGGAGTGTCTGAGGTGCGCCCCACTCTGTACAGAGGCT
GTTTGGGCAGCCTTGCCTCCAGAGAGCAGATTCCAGCTTCGGAAGCCCCTGGTCTAACTTGG
GATCTGGGAATGGAAGGTGCTCCCATCGGAGGGGACCCTCAGAGCCCTGGAGACTGCCAGGT
GGGCCTGCTGCCACTGTAAGCCAAAAGGTGGGGAAGTCCTGACTCCAGGGTCCTTGCCCCAC
CCCTGCCTGCCACCTGGGCCCTCACAGCCCAGACCCTCACTGGGAGGTGAGCTCAGCTGCCC
TTTGGAATAAAGCTGCCTGATCAAAAAAAAAAAAAAAAAAAA

FIGURE 63

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49152
><subunit 1 of 1, 802 aa, 1 stop
><MW: 88846, pI: 6.41, NX(S/T): 7
MPVAEAPQVAGGQGDGGDGEEAEPEGMFKACEDSKRKARGYLRLVPLFVLLALLVLASAGVL
LWYFLGYKAEVMVSQVYSGSLRVLNRHFSQDLTRRESSAFRSETAKAQKMLKELITSTRLGT
YYNSSSVYSFGEGPLTCFFWFILQIPEHRRLMLSPEVVQALLVEELLSTVNSSAAVPYRAEY
EVDPEGLVILEASVKDIAALNSTLGCYRYSYVGQGQVLRLKGPDHLASSCLWHLQGPKDLML
KLRLEWTLAECRDRLAMYDVAGPLEKRLITSVYGCSRQEPVVEVLASGAIMAVVWKKGLHSY
YDPFVLSVQPVVFQACEVNLTLDNRLDSQGVLSTPYFPSYYSPQTHCSWHLTVPSLDYGLAL
WFDAYALRRQKYDLPCTQGQWTIQNRRLCGLRILQPYAERIPVVATAGITINFTSQISLTGP
GVRVHYGLYNQSDPCPGEFLCSVNGLCVPACDGVKDCPNGLDERNCVCRATFQCKEDSTCIS
LPKVCDGQPDCLNGSDEEQCQEGVPCGTFTFQCEDRSCVKKPNPQCDGRPDCRDGSDEEHCD
CGLQGPSSRIVGGAVSSEGEWPWQASLQVRGRHICGGALIADRWVITAAHCFQEDSMASTVL
WTVFLGKVWQNSRWPGEVSFKVSRLLLHPYHEEDSHDYDVALLQLDHPVVRSAAVRPVCLPA
RSHFFEPGLHCWITGWGALREGGPISNALQKVDVQLIPQDLCSEAYRYQVTPRMLCAGYRKG
KKDACQGDSGGPLVCKALSGRWFLAGLVSWGLGCGRPNYFGVYTRITGVISWIQQVVT

Important features:
Type II transmembrane domain:
amino acids 46-67

Serine proteases, trypsin family, histidine active site.
amino acids 604-609

N-glycosylation sites.
amino acids 127-130, 175-178, 207-210, 329-332, 424-427, 444-447 and 509-512

Kringle domains.
amino acids 746-758 and 592-609

Homologous region to Kallikrein Light Chain:
amino acids 568-779

Homologous region to Low-density lipoprotein receptor:
amino acids 451-567

FIGURE 64

GCACCCAGGGCCAGTGGACGATCCAGAACAGGAGGCTGTGTGGCTTGCGCATCCTGCAGCCC
TACGCCGAGAGGATCCCCGTGGTGGCCACGGCCGGGATCACCATCAACTTCACCTCCCAGAT
CTCCCTCACCGGGCCCGGTGTGCGGGTGCACTATGGCTTGTACAACCAGTCGGACCCCTGCC
CTGGAGAGTTCCTCTGTTCTGTGAATGGACTCTGTGTCCCTGCCTGTGATGGGGTCAAGGAC
TGCCCCAACGGCCTGGATGAGAGAAACTGCGTTTGCAGAGCCACATTCCAGTGCAAAGAGGA
CAGCACATGCATCTCACTGCCCAAGGTCTGTGATGGGCAGCCTGATTGTCTCAACGGCAGCG
ATGAAGAGCAGTGCCAGGAAGGGGTGCCATGTGGGACATTCACCTTCCAGTGTGAGGACCGG
AGCTGCGTGAAGAAGCCCAACCCGCAGTGTGATGGGCGGCCCGACTGCAGGGACGGCTCGGA
TGAGGAGCACTGTGACTGTGGCCTCCAGGGCCCCTCCAGCCGCATTGTTGGTGGAGCTGTGT
CCTCCGAGGGTGAGTGGCCATGGCAGGCCAGCCTCCAGGTTCGGGGTCGACACATCTGTGGG
GGGGCCCTCATCGCTGACCGCTGGGTGATAACAGCTGCCCACTGCTTCCAGGAGGACAGCAT
GGCCTCCACGGTGCTGTGGACCGTGTTCCTGGGCAAGGTGTGGCAGAACTCGCGCTGGCCTG
GAGAGGTGTCCTTCAAGGTGAGCCGCCTGCTCCTGCACCCGTACCACGAAGAGGACAGCCAT
GACTACGACGTGGCGCTGCTGCAGCTCGACCACCCGGTGGTGCGCTCGGCCGCCGTGCGCCC
CGTCTGCCTGCCCGCGCGCTCCCACTTCTTCGAGCCCGGCCTGCACTGCTGGATTACGGGCT
GGGGCGCCTTGCGCGAGGGCGGCCCCATCAGCAACGCTCTGCAGAAAGTGGATGTGCAGTTG
ATCCCACAGGACCTGTGCAGCGAGGCCTATCGCTACCAGGTGACGCCACGCATGCTGTGTGC
CGGCTACCGCAAGGGCAAGAAGGATGCCTGTCAGGGTGACTCAGGTGGTCCGCTGGTGTGCA
AGGCACTCAGTGGCCGCTGGTTCCTGGCGGGGCTGGTCAGCTGGGGCCTGGGCTGTGGCCGG
CCTAACTACTTCGGCGTCTACACCCGCATCACAGGTGTGATCAGCTGGATCCAGCAAGTGGT
GACCTGAGGAACTGCCCCCCTGCAAAGCAGGGCCCACCTCCTGGACTCAGAGAGCCCAGGGC
AACTGCCAAGCAGGGGGACAAGTAT

FIGURE 65

GGACGAGGGCAGATCTCGTTCTGGGGCAAGCCGTTGACACTCGCTCCCTGCCACCGCCCGGG
CTCCGTGCCGCCAAGTTTTCATTTTCCACCTTCTCTGCCTCCAGTCCCCCAGCCCCTGGCCG
AGAGAAGGGTCTTACCGGCCGGGATTGCTGGAAACACCAAGAGGTGGTTTTTGTTTTTTAAA
ACTTCTGTTTCTTGGGAGGGGGTGTGGCGGGGCAGGATGAGCAACTCCGTTCCTCTGCTCTG
TTTCTGGAGCCTCTGCTATTGCTTTGCTGCGGGGAGCCCCGTACCTTTTGGTCCAGAGGGAC
GGCTGGAAGATAAGCTCCACAAACCCAAAGCTACACAGACTGAGGTCAAACCATCTGTGAGG
TTTAACCTCCGCACCTCCAAGGACCCAGAGCATGAAGGATGCTACCTCTCCGTCGGCCACAG
CCAGCCCTTAGAAGACTGCAGTTTCAACATGACAGCTAAAACCTTTTTCATCATTCACGGAT
GGACGATGAGCGGTATCTTTGAAAACTGGCTGCACAAACTCGTGTCAGCCCTGCACACAAGA
GAGAAAGACGCCAATGTAGTTGTGGTTGACTGGCTCCCCCTGGCCCACCAGCTTTACACGGA
TGCGGTCAATAATACCAGGGTGGTGGGACACAGCATTGCCAGGATGCTCGACTGGCTGCAGG
AGAAGGACGATTTTTCTCTCGGGAATGTCCACTTGATCGGCTACAGCCTCGGAGCGCACGTG
GCCGGGTATGCAGGCAACTTCGTGAAAGGAACGGTGGGCCGAATCACAGGTTTGGATCCTGC
CGGGCCCATGTTTGAAGGGGCCGACATCCACAAGAGGCTCTCTCCGGACGATGCAGATTTTG
TGGATGTCCTCCACACCTACACGCGTTCCTTCGGCTTGAGCATTGGTATTCAGATGCCTGTG
GGCCACATTGACATCTACCCCAATGGGGGTGACTTCCAGCCAGGCTGTGGACTCAACGATGT
CTTGGGATCAATTGCATATGGAACAATCACAGAGGTGGTAAAATGTGAGCATGAGCGAGCCG
TCCACCTCTTTGTTGACTCTCTGGTGAATCAGGACAAGCCGAGTTTTGCCTTCCAGTGCACT
GACTCCAATCGCTTCAAAAAGGGGATCTGTCTGAGCTGCCGCAAGAACCGTTGTAATAGCAT
TGGCTACAATGCCAAGAAAATGAGGAACAAGAGGAACAGCAAAATGTACCTAAAAACCCGGG
CAGGCATGCCTTTCAGAGGTAACCTTCAGTCCCTGGAGTGTCCCTGAGGAAGGCCCTTAATA
CCTCCTTCTTAATACCATGCTGCAGAGCAGGGCACATCCTAGCCCAGGAGAAGTGGCCAGCA
CAATCCAATCAAATCGTTGCAAATCAGATTACACTGTGCATGTCCTAGGAAAGGGAATCTTT
ACAAAATAAACAGTGTGGACCCCTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAA

FIGURE 66

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49646
><subunit 1 of 1, 354 aa, 1 stop
><MW: 39362, pI: 8.35, NX(S/T): 2
MSNSVPLLCFWSLCYCFAAGSPVPFGPEGRLEDKLHKPKATQTEVKPSVRFNLRTSKDPEHE
GCYLSVGHSQPLEDCSFNMTAKTFFIIHGWTMSGIFENWLHKLVSALHTREKDANVVVVDWL
PLAHQLYTDAVNNTRVVGHSIARMLDWLQEKDDFSLGNVHLIGYSLGAHVAGYAGNFVKGTV
GRITGLDPAGPMFEGADIHKRLSPDDADFVDVLHTYTRSFGLSIGIQMPVGHIDIYPNGGDF
QPGCGLNDVLGSIAYGTITEVVKCEHERAVHLFVDSLVNQDKPSFAFQCTDSNRFKKGICLS
CRKNRCNSIGYNAKKMRNKRNSKMYLKTRAGMPFRGNLQSLECP

Important features:
Signal peptide:
amino acids 1-16

Lipases, serine active site.
amino acids 163-172

N-glycosylation sites.
amino acids 80-83 and 136-139

FIGURE 67

CGGACGCGTGGGCGGACGCGTGGGCCTGGGCAAGGGCCGGGGCGCCGGGCCGAGCCACCTCT
TCCCCTCCCCCGCTTCCCTGTCGCGCTCCGCTGGCTGGACGCGCTGGAGGAGTGGAGCAGCA
CCCGGCCGGCCCTGGGGGCTGACAGTCGGCAAAGTTTGGCCCGAAGAGGAAGTGGTCTCAAA
CCCCGGCAGGTGGCGACCAGGCCAGACCAGGGGCGCTCGCTGCCTGCGGGCGGGCTGTAGGC
GAGGGCGCGCCCCAGTGCCGAGACCCGGGGCTTCAGGAGCCGGCCCCGGGAGAGAAGAGTGC
GGCGGCGGACGGAGAAAACAACTCCAAAGTTGGCGAAAGGCACCGCCCCTACTCCCGGGCTG
CCGCCGCCTCCCCGCCCCCAGCCCTGGCATCCAGAGTACGGGTCGAGCCCGGGCCATGGAGC
CCCCCTGGGGAGGCGGCACCAGGGAGCCTGGGCGCCCGGGGCTCCGCCGCGACCCCATCGGG
TAGACCACAGAAGCTCCGGGACCCTTCCGGCACCTCTGGACAGCCCAGGATGCTGTTGGCCA
CCCTCCTCCTCCTCCTCCTTGGAGGCGCTCTGGCCCATCCAGACCGGATTATTTTTCCAAAT
CATGCTTGTGAGGACCCCCCAGCAGTGCTCTTAGAAGTGCAGGGCACCTTACAGAGGCCCCT
GGTCCGGGACAGCCGCACCTCCCCTGCCAACTGCACCTGGCTCATCCTGGGCAGCAAGGAAC
AGACTGTCACCATCAGGTTCCAGAAGCTACACCTGGCCTGTGGCTCAGAGCGCTTAACCCTA
CGCTCCCCTCTCCAGCCACTGATCTCCCTGTGTGAGGCACCTCCCAGCCCTCTGCAGCTGCC
CGGGGGCAACGTCACCATCACTTACAGCTATGCTGGGGCCAGAGCACCCATGGGCCAGGGCT
TCCTGCTCTCCTACAGCCAAGATTGGCTGATGTGCCTGCAGGAAGAGTTTCAGTGCCTGAAC
CACCGCTGTGTATCTGCTGTCCAGCGCTGTGATGGGGTTGATGCCTGTGGCGATGGCTCTGA
TGAAGCAGGTTGCAGCTCAGACCCCTTCCCTGGCCTGACCCCAAGACCCGTCCCCTCCCTGC
CTTGCAATGTCACCTTGGAGGACTTCTATGGGGTCTTCTCCTCTCCTGGATATACACACCTA
GCCTCAGTCTCCCACCCCCAGTCCTGCCATTGGCTGCTGGACCCCCATGATGGCCGGCGGCT
GGCCGTGCGCTTCACAGCCCTGGACTTGGGCTTTGGAGATGCAGTGCATGTGTATGACGGCC
CTGGGCCCCCTGAGAGCTCCCGACTACTGCGTAGTCTCACCCACTTCAGCAATGGCAAGGCT
GTCACTGTGGAGACACTGTCTGGCCAGGCTGTTGTGTCCTACCACACAGTTGCTTGGAGCAA
TGGTCGTGGCTTCAATGCCACCTACCATGTGCGGGGCTATTGCTTGCCTTGGGACAGACCCT
GTGGCTTAGGCTCTGGCCTGGGAGCTGGCGAAGGCCTAGGTGAGCGCTGCTACAGTGAGGCA
CAGCGCTGTGACGGCTCATGGGACTGTGCTGACGGCACAGATGAGGAGGACTGCCCAGGCTG
CCCACCTGGACACTTCCCCTGTGGGGCTGCTGGCACCTCTGGTGCCACAGCCTGCTACCTGC
CTGCTGACCGCTGCAACTACCAGACTTTCTGTGCTGATGGAGCAGATGAGAGACGCTGTCGG
CATTGCCAGCCTGGCAATTTCCGATGCCGGGACGAGAAGTGCGTGTATGAGACGTGGGTGTG
CGATGGGCAGCCAGACTGTGCGGACGGCAGTGATGAGTGGGACTGCTCCTATGTTCTGCCCC
GCAAGGTCATTACAGCTGCAGTCATTGGCAGCCTAGTGTGCGGCCTGCTCCTGGTCATCGCC
CTGGGCTGCACCTGCAAGCTCTATGCCATTCGCACCCAGGAGTACAGCATCTTTGCCCCCCT
CTCCCGGATGGAGGCTGAGATTGTGCAGCAGCAGGCACCCCCTTCCTACGGGCAGCTCATTG
CCCAGGGTGCCATCCCACCTGTAGAAGACTTTCCTACAGAGAATCCTAATGATAACTCAGTG
CTGGGCAACCTGCGTTCTCTGCTACAGATCTTACGCCAGGATATGACTCCAGGAGGTGGCCC
AGGTGCCCGCCGTCGTCAGCGGGGCCGCTTGATGCGACGCCTGGTACGCCGTCTCCGCCGCT
GGGGCTTGCTCCCTCGAACCAACACCCCGGCTCGGGCCTCTGAGGCCAGATCCCAGGTCACA
CCTTCTGCTGCTCCCCTTGAGGCCCTAGATGGTGGCACAGGTCCAGCCCGTGAGGGCGGGGC
AGTGGGTGGGCAAGATGGGGAGCAGGCACCCCCACTGCCCATCAAGGCTCCCCTCCCATCTG
CTAGCACGTCTCCAGCCCCCACTACTGTCCCTGAAGCCCCAGGGCCACTGCCCTCACTGCCC
CTAGAGCCATCACTATTGTCTGGAGTGGTGCAGGCCCTGCGAGGCCGCCTGTTGCCCAGCCT
GGGGCCCCCAGGACCAACCCGGAGCCCCCCTGGACCCCACACAGCAGTCCTGGCCCTGGAAG
ATGAGGACGATGTGCTACTGGTGCCACTGGCTGAGCCGGGGGTGTGGGTAGCTGAGGCAGAG
GATGAGCCACTGCTTACCTGAGGGGACCTGGGGGCTCTACTGAGGCCTCTCCCCTGGGGGCT
CTACTCATAGTGGCACAACCTTTTAGAGGTGGGTCAGCCTCCCCTCCACCACTTCCTTCCCT
GTCCCTGGATTTCAGGGACTTGGTGGGCCTCCCGTTGACCCTATGTAGCTGCTATAAAGTTA
AGTGTCCCTCAGGCAGGGAGAGGGCTCACAGAGTCTCCTCTGTACGTGGCCATGGCCAGACA
CCCCAGTCCCTTCACCACCACCTGCTCCCCACGCCACCACCATTTGGGTGGCTGTTTTTAAA
AAGTAAAGTTCTTAGAGGATCATAGGTCTGGACACTCCATCCTTGCCAAACCTCTACCCAAA
AGTGGCCTTAAGCACCGGAATGCCAATTAACTAGAGACCCTCCAGCCCCCAAGGGGAGGATT
TGGGCAGAACCTGAGGTTTTGCCATCCACAATCCCTCCTACAGGGCCTGGCTCACAAAAAGA
GTGCAACAAATGCTTCTATTCCATAGCTACGGCATTGCTCAGTAAGTTGAGGTCAAAAATAA
AGGAATCATACATCTC

FIGURE 68

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49631
<subunit 1 of 1, 713 aa, 1 stop
<MW: 76193, pI: 5.42, NX(S/T): 4
MLLATLLLLLLGGALAHPDRIIFPNHACEDPPAVLLEVQGTLQRPLVRDSRTSPANCTWLIL
GSKEQTVTIRFQKLHLACGSERLTLRSPLQPLISLCEAPPSPLQLPGGNVTITYSYAGARAP
MGQGFLLSYSQDWLMCLQEEFQCLNHRCVSAVQRCDGVDACGDGSDEAGCSSDPFPGLTPRP
VPSLPCNVTLEDFYGVFSSPGYTHLASVSHPQSCHWLLDPHDGRRLAVRFTALDLGFGDAVH
VYDGPGPPESSRLLRSLTHFSNGKAVTVETLSGQAVVSYHTVAWSNGRGFNATYHVRGYCLP
WDRPCGLGSGLGAGEGLGERCYSEAQRCDGSWDCADGTDEEDCPGCPPGHFPCGAAGTSGAT
ACYLPADRCNYQTFCADGADERRCRHCQPGNFRCRDEKCVYETWVCDGQPDCADGSDEWDCS
YVLPRKVITAAVIGSLVCGLLLVIALGCTCKLYAIRTQEYSIFAPLSRMEAEIVQQQAPPSY
GQLIAQGAIPPVEDFPTENPNDNSVLGNLRSLLQILRQDMTPGGGPGARRRQRGRLMRRLVR
RLRRWGLLPRTNTPARASEARSQVTPSAAPLEALDGGTGPAREGGAVGGQDGEQAPPLPIKA
PLPSASTSPAPTTVPEAPGPLPSLPLEPSLLSGVVQALRGRLLPSLGPPGPTRSPPGPHTAV
LALEDEDDVLLVPLAEPGVWVAEAEDEPLLT

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 442-462

LDL-receptor class A (LDLRA) domain proteins

amino acids 411-431, 152-171, 331-350 and 374-393

FIGURE 69

CGAGCTGGGCGAGAAGTAGGGGAGGGCGGTGCTCCGCCGCGGTGGCGGTTGCTATCGCTTCG
CAGAACCTACTCAGGCAGCCAGCTGAGAAGAGTTGAGGGAAAGTGCTGCTGCTGGGTCTGCA
GACGCGATGGATAACGTGCAGCCGAAAATAAAACATCGCCCCTTCTGCTTCAGTGTGAAAGG
CCACGTGAAGATGCTGCGGCTGGCACTAACTGTGACATCTATGACCTTTTTTATCATCGCAC
AAGCCCCTGAACCATATATTGTTATCACTGGATTTGAAGTCACCGTTATCTTATTTTTCATA
CTTTTATATGTACTCAGACTTGATCGATTAATGAAGTGGTTATTTTGGCCTTTGCTTGATAT
TATCAACTCACTGGTAACAACAGTATTCATGCTCATCGTATCTGTGTTGGCACTGATACCAG
AAACCACAACATTGACAGTTGGTGGAGGGGTGTTTGCACTTGTGACAGCAGTATGCTGTCTT
GCCGACGGGGCCCTTATTTACCGGAAGCTTCTGTTCAATCCCAGCGGTCCTTACCAGAAAAA
GCCTGTGCATGAAAAAAAAGAAGTTTTGTAATTTTATATTACTTTTTAGTTTGATACTAAGT
ATTAAACATATTTCTGTATTCTTCCAAAAAAAAAAAAAAAAAA

FIGURE 70

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49645
><subunit 1 of 1, 152 aa, 1 stop
><MW: 17170, pI: 9.62, NX(S/T): 1
MDNVQPKIKHRPFCFSVKGHVKMLRLALTVTSMTFFIIAQAPEPYIVITGFEVTVILFFILL
YVLRLDRLMKWLFWPLLDIINSLVTTVFMLIVSVLALIPETTTLTVGGGVFALVTAVCCLAD
GALIYRKLLFNPSGPYQKKPVHEKKEVL

Important features:

Potential type II transmembrane domain:
amino acids 26-42

Other potential transmembrane domain:
amino acids 44-65, 81-101 and 109-129

Leucine zipper pattern
amino acids 78-99 and 85-106

N-myristoylation site.
amino acids 110-115

Ribonucleotide reductase large subunit protein
amino acids 116-127

FIGURE 71

GGGCGAGAAGTAGGGGAGGGCGTGTTCCGCCGCGGTGGCGGTTGCTATCGTTTTGCAGAACC
TACTCAGGCAGCCAGNTGAGAAGAGTTGAGGGAAAGTGCTGCTGCTGGGTCTGCAGACGCGA
TGGATAACGTGCAGCCGAAAATAAAACATCGCCCCTTCTGCTTCAGTGTGAAAGGCCACGTG
AAGATGCTGCGGCTGGCACTAACTGNGACATCTATGACCTTTTTTATNATCGCACAAGCCCC
TGAACCATATATTGTTATCACTGGATTTGAAGTCACCGTTATCTTATTTTTCATACTTTTAT
ATGTACTCAGACTTGATCGATTAATGAAGTGGTTATTTTGGCCTTTGCTTGATATTATCAAC
TCACTGGTAACAACAGTATTCATGCTCATCGTATCTGTGTTGGCACTGATACCAGAAACCAC
AACATTGACAGTTGGTGGAGGGGTGTTTGCACTTGTGACAGCAGTATGCTGTNTTGCCGAC

FIGURE 72

CAGCCCCGCGCGCCGGCCGAGTCGCTGAGCCGCGGCTGCCGGACGGGACGGGACCGGCTAGG
CTGGGCGCGCCCCCCGGGCCCCGCCGTGGGCATGGGCGCACTGGCCCGGGCGCTGCTGCTGC
CTCTGCTGGCCCAGTGGCTCCTGCGCGCCGCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTG
CCCCTCCGGGTGGCCGCGGCCACGAACCGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCC
TGCCGAGCGCCACGCCGACGGCTTGGCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGG
GCGCCGCCAACTTCTTGGCCATGGTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTAC
CTGGAGATGCTGATCGGGACCCCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAG
TAACTTTGCCGTGGCAGGAACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGT
CTAGCACATACCGCTCCAAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACG
GGCTTCGTTGGGGAAGACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAA
CATTGCCACTATTTTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATAC
TTGGCCTAGCTTATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCC
CTGGTGACACAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCC
CGTTGCTGGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGT
ATAAAGGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTG
AAATTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGC
CATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTGG
AAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCCCAG
CTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATCTACCT
GAGAGACGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTACATTCAGC
CCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCATCCACAAAT
GCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGACAGAGCCCAGAA
GAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCAGTGTCTGAAATTT
CCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCTCAGTCTTTGAGCGAG
CCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGAGCCATCCTCCTTGTCTT
AATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCCCGTGACCCTGAGGTCGTCA
ATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAATGAATAGCCAGGCCTGACCTCAAGCAA
CCATGAACTCAGCTATTAAGAAAATCACATTTCCAGGGCAGCAGCCGGGATCGATGGTGGCG
CTTTCTCCTGTGCCCACCCGTCTTCAATCTCTGTTCTGCTCCCAGATGCCTTCTAGATTCAC
TGTCTTTTGATTCTTGATTTTCAAGCTTTCAAATCCTCCCTACTTCCAAGAAAAATAATTAA
AAAAAAAACTTCATTCTAA

FIGURE 73

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45493
><subunit 1 of 1, 518 aa, 1 stop
><MW: 56180, pI: 5.08, NX(S/T): 2
MGALARALLLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGLAL
ALEPALASPAGAANFLAMVDNLQGDSGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAGTPHS
YIDTYFDTERSSTYRSKGFDVTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATIFESENF
FLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVAGSGTNGGS
LVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLDCREYNADKAIVDSGTTLLR
LPQKVFDAVVEAVARASLIPEFSDGFWTGSQLACWTNSETPWSYFPKISIYLRDENSSRSFR
ITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFDRAQKRVGFAASPCA
EIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCGAILLVLIVLLLLPFRC
QRRPRDPEVVNDESSLVRHRWK

Important features:

Signal peptide:
amino acids 1-20

Transmembrane domain:
amino acids 466-494

N-glycosylation sites.
amino acids 170-173 and 366-369

Leucine zipper pattern.
amino acids 10-31 and 197-118

Eukaryotic and viral aspartyl proteases
amino acids 109-118, 252-261 and 298-310

FIGURE 74

CGCCTCCGCCTTCGGAGGCTGACGCGCCCGGGCGCCGTTCCAGGCCTGTGCAGGGCGGATCG
GCAGCCGCCTGGCGGCGATCCAGGGCGGTGCGGGGCCTGGGCGGGAGCCGGGAGGCGCGGCC
GGCATGGAGGCGCTGCTGCTGGGCGCGGGGTTGCTGCTGGGCGCTTACGTGCTTGTCTACTA
CAACCTGGTGAAGGCCCCGCCGTGCGGCGGCATGGGCAACCTGCGGGGCCGCACGGCCGTGG
TCACGGGCGCCAACAGCGGCATCGGAAAGATGACGGCGCTGGAGCTGGCGCGCCGGGGAGCG
CGCGTGGTGCTGGCCTGCCGCAGCCAGGAGCGCGGGGAGGCGGCTGCCTTCGACCTCCGCCA
GGAGAGTGGGAACAATGAGGTCATCTTCATGGCCTTGGACTTGGCCAGTCTGGCCTCGGTGC
GGGCCTTTGCCACTGCCTTTCTGAGCTCTGAGCCACGGTTGGACATCCTCATCCACAATGCC
GGTATCAGTTCCTGTGGCCGGACCCGTGAGGCGTTTAACCTGCTGCTTCGGGTGAACCATAT
CGGTCCCTTTCTGCTGACACATCTGCTGCTGCCTTGCCTGAAGGCATGTGCCCCTAGCCGCG
TGGTGGTGGTAGCCTCAGCTGCCCACTGTCGGGGACGTCTTGACTTCAAACGCCTGGACCGC
CCAGTGGTGGGCTGGCGGCAGGAGCTGCGGGCATATGCTGACACTAAGCTGGCTAATGTACT
GTTTGCCCGGGAGCTCGCCAACCAGCTTGAGGCCACTGGCGTCACCTGCTATGCAGCCCACC
CAGGGCCTGTGAACTCGGAGCTGTTCCTGCGCCATGTTCCTGGATGGCTGCGCCCACTTTTG
CGCCCATTGGCTTGGCTGGTGCTCCGGGCACCAAGAGGGGGTGCCCAGACACCCCTGTATTG
TGCTCTACAAGAGGGCATCGAGCCCCTCAGTGGGAGATATTTTGCCAACTGCCATGTGGAAG
AGGTGCCTCCAGCTGCCCGAGACGACCGGGCAGCCCATCGGCTATGGGAGGCCAGCAAGAGG
CTGGCAGGGCTTGGGCCTGGGGAGGATGCTGAACCCGATGAAGACCCCCAGTCTGAGGACTC
AGAGGCCCCATCTTCTCTAAGCACCCCCCACCCTGAGGAGCCCACAGTTTCTCAACCTTACC
CCAGCCCTCAGAGCTCACCAGATTTGTCTAAGATGACGCACCGAATTCAGGCTAAAGTTGAG
CCTGAGATCCAGCTCTCCTAACCCTCAGGCCAGGATGCTTGCCATGGCACTTCATGGTCCTT
GAAAACCTCGGATGTGTGTGAGGCCATGCCCTGGACACTGACGGGTTTGTGATCTTGACCTC
CGTGGTTACTTTCTGGGGCCCCAAGCTGTGCCCTGGACATCTCTTTTCCTGGTTGAAGGAAT
AATGGGTGATTATTTCTTCCTGAGAGTGACAGTAACCCCAGATGGAGAGATAGGGGTATGCT
AGACACTGTGCTTCTCGGAAATTTGGATGTAGTATTTTCAGGCCCCACCCTTATTGATTCTG
ATCAGCTCTGGAGCAGAGGCAGGGAGTTTGCAATGTGATGCACTGCCAACATTGAGAATTAG
TGAACTGATCCCTTTGCAACCGTCTAGCTAGGTAGTTAAATTACCCCCATGTTAATGAAGCG
GAATTAGGCTCCCGAGCTAAGGGACTCGCCTAGGGTCTCACAGTGAGTAGGAGGAGGGCCTG
GGATCTGAACCCAAGGGTCTGAGGCCAGGGCCGACTGCCGTAAGATGGGTGCTGAGAAGTGA
GTCAGGGCAGGGCAGCTGGTATCGAGGTGCCCCATGGGAGTAAGGGGACGCCTTCCGGGCGG
ATGCAGGGCTGGGGTCATCTGTATCTGAAGCCCCTCGGAATAAAGCGCGTTGACCGCCAAAA
AAAAAAAAAAAAAAAA

FIGURE 75

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48227

<subunit 1 of 1, 377 aa, 1 stop

<MW: 40849, pI: 7.98, NX(S/T): 0

MEALLLGAGLLLGAYVLVYYNLVKAPPCGGMGNLRGRTAVVTGANSGIGKMTALELARRGAR
VVLACRSQERGEAAAFDLRQESGNNEVIFMALDLASLASVRAFATAFLSSEPRLDILIHNAG
ISSCGRTREAFNLLLRVNHIGPFLLTHLLLPCLKACAPSRVVVVASAAHCRGRLDFKRLDRP
VVGWRQELRAYADTKLANVLFARELANQLEATGVTCYAAHPGPVNSELFLRHVPGWLRPLLR
PLAWLVLRAPRGGAQTPLYCALQEGIEPLSGRYFANCHVEEVPPAARDDRAAHRLWEASKRL
AGLGPGEDAEPDEDPQSEDSEAPSSLSTPHPEEPTVSQPYPSPQSSPDLSKMTHRIQAKVEP
EIQLS

Important features:

Signal peptide:

amino acids 1-16

Glycosaminoglycan attachment site.

amino acids 46-49

Short-chain alcohol dehydrogenase family

amino acids 37-49 and 114-124

FIGURE 76A

GGAGGAGACAGCCTCCTGGGGGGCAGGGGTTCCCTGCCTCTGCTGCTCCTGCTCATCATGGG
AGGCATGGCTCAGGACTCCCCGCCCCAGATCCTAGTCCACCCCCAGGACCAGCTGTTCCAGG
GCCCTGGCCCTGCCAGGATGAGCTGCCAAGCCTCAGGCCAGCCACCTCCCACCATCCGCTGG
TTGCTGAATGGGCAGCCCCTGAGCATGGTGCCCCCAGACCCACACCACCTCCTGCCTGATGG
GACCCTTCTGCTGCTACAGCCCCCTGCCCGGGGACATGCCCACGATGGCCAGGCCCTGTCCA
CAGACCTGGGTGTCTACACATGTGAGGCCAGCAACCGGCTTGGCACGGCAGTCAGCAGAGGC
GCTCGGCTGTCTGTGGCTGTCCTCCGGGAGGATTTCCAGATCCAGCCTCGGGACATGGTGGC
TGTGGTGGGTGAGCAGTTTACTCTGGAATGTGGGCCGCCCTGGGGCCACCCAGAGCCCACAG
TCTCATGGTGGAAAGATGGGAAACCCCTGGCCCTCCAGCCCGGAAGGCACACAGTGTCCGGG
GGGTCCCTGCTGATGGCAAGAGCAGAGAAGAGTGACGAAGGGACCTACATGTGTGTGGCCAC
CAACAGCGCAGGACATAGGGAGAGCCGCGCAGCCCGGGTTTCCATCCAGGAGCCCCAGGACT
ACACGGAGCCTGTGGAGCTTCTGGCTGTGCGAATTCAGCTGGAAAATGTGACACTGCTGAAC
CCGGATCCTGCAGAGGGCCCCAAGCCTAGACCGGCGGTGTGGCTCAGCTGGAAGGTCAGTGG
CCCTGCTGCGCCTGCCCAATCTTACACGGCCTTGTTCAGGACCCAGACTGCCCCGGGAGGCC
AGGGAGCTCCGTGGGCAGAGGAGCTGCTGGCCGGCTGGCAGAGCGCAGAGCTTGGAGGCCTC
CACTGGGGCCAAGACTACGAGTTCAAAGTGAGACCATCCTCTGGCCGGGCTCGAGGCCCTGA
CAGCAACGTGCTGCTCCTGAGGCTGCCGGAAAAAGTGCCCAGTGCCCCACCTCAGGAAGTGA
CTCTAAAGCCTGGCAATGGCACTGTCTTTGTGAGCTGGGTCCCACCACCTGCTGAAAACCAC
AATGGCATCATCCGTGGCTACCAGGTCTGGAGCCTGGGCAACACATCACTGCCACCAGCCAA
CTGGACTGTAGTTGGTGAGCAGACCCAGCTGGAAATCGCCACCCATATGCCAGGCTCCTACT
GCGTGCAAGTGGCTGCAGTCACTGGTGCTGGAGCTGGGGAGCCCAGTAGACCTGTCTGCCTC
CTTTTAGAGCAGGCCATGGAGCGAGCCACCCAAGAACCCAGTGAGCATGGTCCCTGGACCCT
GGAGCAGCTGAGGGCTACCTTGAAGCGGCCTGAGGTCATTGCCACCTGCGGTGTTGCACTCT
GGCTGCTGCTTCTGGGCACCGCCGTGTGTATCCACCGCCGGCGCCGAGCTAGGGTGCACCTG
GGCCCAGGTCTGTACAGATATACCAGTGAGGATGCCATCCTAAAACACAGGATGGATCACAG
TGACTCCCAGTGGTTGGCAGACACTTGGCGTTCCACCTCTGGCTCTCGGGACCTGAGCAGCA
GCAGCAGCCTCAGCAGTCGGCTGGGGGCGGATGCCCGGGACCCACTAGACTGTCGTCGCTCC
TTGCTCTCCTGGGACTCCCGAAGCCCCGGCGTGCCCCTGCTTCCAGACACCAGCACTTTTTA
TGGCTCCCTCATCGCTGAGCTGCCCTCCAGTACCCCAGCCAGGCCAAGTCCCCAGGTCCCAG
CTGTCAGGCGCCTCCCACCCCAGCTGGCCCAGCTCTCCAGCCCCTGTTCCAGCTCAGACAGC
CTCTGCAGCCGCAGGGGACTCTCTTCTCCCCGCTTGTCTCTGGCCCCTGCAGAGGCTTGGAA
GGCCAAAAAGAAGCAGGAGCTGCAGCATGCCAACAGTTCCCCACTGCTCCGGGGCAGCCACT
CCTTGGAGCTCCGGGCCTGTGAGTTAGGAAATAGAGGTTCCAAGAACCTTTCCCAAAGCCCA
GGAGCTGTGCCCCAAGCTCTGGTTGCCTGGCGGGCCCTGGGACCGAAACTCCTCAGCTCCTC
AAATGAGCTGGTTACTCGTCATCTCCCTCCAGCACCCCTCTTTCCTCATGAAACTCCCCCAA
CTCAGAGTCAACAGACCCAGCCTCCGGTGGCACCACAGGCTCCCTCCTCCATCCTGCTGCCA
GCAGCCCCCATCCCCATCCTTAGCCCCTGCAGTCCCCCTAGCCCCCAGGCCTCTTCCCTCTC
TGGCCCCAGCCCAGCTTCCAGTCGCCTGTCCAGCTCCTCACTGTCATCCCTGGGGGAGGATC
AAGACAGCGTGCTGACCCCTGAGGAGGTAGCCCTGTGCTTGGAACTCAGTGAGGGTGAGGAG
ACTCCCAGGAACAGCGTCTCTCCCATGCCAAGGGCTCCTTCACCCCCCACCACCTATGGGTA
CATCAGCGTCCCAACAGCCTCAGAGTTCACGGACATGGGCAGGACTGGAGGAGGGGTGGGGC
CCAAGGGGGGAGTCTTGCTGTGCCCACCTCGGCCCTGCCTCACCCCCACCCCCAGCGAGGGC
TCCTTAGCCAATGGTTGGGGCTCAGCCTCTGAGGACAATGCCGCCAGCGCCAGAGCCAGCCT
TGTCAGCTCCTCCGATGGCTCCTTCCTCGCTGATGCTCACTTTGCCCGGGCCCTGGCAGTGG
CTGTGGATAGCTTTGGTTTCGGTCTAGAGCCCAGGGAGGCAGACTGCGTCTTCATAGATGCC
TCATCACCTCCCTCCCCACGGGATGAGATCTTCCTGACCCCCAACCTCTCCCTGCCCCTGTG
GGAGTGGAGGCCAGACTGGTTGGAAGACATGGAGGTCAGCCACACCCAGCGGCTGGGAAGGG
GGATGCCTCCCTGGCCCCCTGACTCTCAGATCTCTTCCCAGAGAAGTCAGCTCCACTGTCGT
ATGCCCAAGGCTGGTGCTTCTCCTGTAGATTACTCCTGAACCGTGTCCCTGAGACTTCCCAG
ACGGGAATCAGAACCACTTCTCCTGTCCACCCACAAGACCTGGGCTGTGGTGTGTGGGTCTT
GGCCTGTGTTTCTCTGCAGCTGGGGTCCACCTTCCCAAGCCTCCAGAGAGTTCTCCCTCCAC
GATTGTGAAAACAAATGAAAACAAAATTAGAGCAAAGCTGACCTGGAGCCCTCAGGGAGCAA
AACATCATCTCCACCTGACTCCTAGCCACTGCTTTCTCCTCTGTGCCATCCACTCCCACCAC
CAGGTTGTTTTGGCCTGAGGAGCAGCCCTGCCTGCTGCTCTTCCCCCACCATTTGGATCACA

FIGURE 76B

GGAAGTGGAGGAGCCAGAGGTGCCTTTGTGGAGGACAGCAGTGGCTGCTGGGAGAGGGCTGT
GGAGGAAGGAGCTTCTCGGAGCCCCCTCTCAGCCTTACCTGGGCCCCTCCTCTAGAGAAGAG
CTCAACTCTCTCCCAACCTCACCATGGAAAGAAAATAATTATGAATGCCACTGAGGCACTGA
GGCCCTACCTCATGCCAAACAAAGGGTTCAAGGCTGGGTCTAGCGAGGATGCTGAAGGAAGG
GAGGTATGAGACCGTAGGTCAAAAGCACCATCCTCGTACTGTTGTCACTATGAGCTTAAGAA
ATTTGATACCATAAAATGGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 77

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA41404
<subunit 1 of 1, 985 aa, 1 stop
<MW: 105336, pI: 6.55, NX(S/T): 7
MGGMAQDSPPQILVHPQDQLFQGPGPARMSCQASGQPPPTIRWLLNGQPLSMVPPDPHHLLP
DGTLLLLQPPARGHAHDGQALSTDLGVYTCEASNRLGTAVSRGARLSVAVLREDFQIQPRDM
VAVVGEQFTLECGPPWGHPEPTVSWWKDGKPLALQPGRHTVSGGSLLMARAEKSDEGTYMCV
ATNSAGHRESRAARVSIQEPQDYTEPVELLAVRIQLENVTLLNPDPAEGPKPRPAVWLSWKV
SGPAAPAQSYTALFRTQTAPGGQGAPWAEELLAGWQSAELGGLHWGQDYEFKVRPSSGRARG
PDSNVLLLRLPEKVPSAPPQEVTLKPGNGTVFVSWVPPPAENHNGIIRGYQVWSLGNTSLPP
ANWTVVGEQTQLEIATHMPGSYCVQVAAVTGAGAGEPSRPVCLLLEQAMERATQEPSEHGPW
TLEQLRATLKRPEVIATCGVALWLLLLGTAVCIHRRRRARVHLGPGLYRYTSEDAILKHRMD
HSDSQWLADTWRSTSGSRDLSSSSSLSSRLGADARDPLDCRRSLLSWDSRSPGVPLLPDTST
FYGSLIAELPSSTPARPSPQVPAVRRLPPQLAQLSSPCSSSDSLCSRRGLSSPRLSLAPAEA
WKAKKKQELQHANSSPLLRGSHSLELRACELGNRGSKNLSQSPGAVPQALVAWRALGPKLLS
SSNELVTRHLPPAPLFPHETPPTQSQQTQPPVAPQAPSSILLPAAPIPILSPCSPPSPQASS
LSGPSPASSRLSSSSLSSLGEDQDSVLTPEEVALCLELSEGEETPRNSVSPMPRAPSPPTTY
GYISVPTASEFTDMGRTGGGVGPKGGVLLCPPRPCLTPTPSEGSLANGWGSASEDNAASARA
SLVSSSDGSFLADAHFARALAVAVDSFGFGLEPREADCVFIDASSPPSPRDEIFLTPNLSLP
LWEWRPDWLEDMEVSHTQRLGRGMPPWPPDSQISSQRSQLHCRMPKAGASPVDYS

Important features:

Transmembrane domain:

amino acids 448-467

N-glycosylation sites:

amino acids 224-227, 338-341, 367-370, 374-377, 658-661 and 926-929

N-myristoylation sites.

amino acids 47-52, 80-85, 88-93, 99-104, 105-110, 181-186, 272-277, 290-295, 355-360, 403-408, 462-467, 561-566, 652-657, 849-854 and 876-881

Phosphotyrosine interaction domain proteins

amino acids 740-753

FIGURE 78

CTCCCACGGTGTCCAGCGCCCAGAATGCGGCTTCTGGTCCTGCTATGGGGTTGCCTGCTGCT
CCCAGGTTATGAAGCCCTGGAGGGCCCAGAGGAAATCAGCGGGTTCGAAGGGGACACTGTGT
CCCTGCAGTGCACCTACAGGGAAGAGCTGAGGGACCACCGGAAGTACTGGTGCAGGAAGGGT
GGGATCCTCTTCTCTCGCTGCTCTGGCACCATCTATGCAGAAGAAGAAGGCCAGGAGACAAT
GAAGGGCAGGGTGTCCATCCGTGACAGCCGCCAGGAGCTCTCGCTCATTGTGACCCTGTGGA
ACCTCACCCTGCAAGACGCTGGGGAGTACTGGTGTGGGGTCGAAAAACGGGGCCCCGATGAG
TCTTTACTGATCTCTCTGTTCGTCTTTCCAGGACCCTGCTGTCCTCCCTCCCCTTCTCCCAC
CTTCCAGCCTCTGGCTACAACACGCCTGCAGCCCAAGGCAAAAGCTCAGCAAACCCAGCCCC
CAGGATTGACTTCTCCTGGGCTCTACCCGGCAGCCACCACAGCCAAGCAGGGGAAGACAGGG
GCTGAGGCCCCTCCATTGCCAGGGACTTCCCAGTACGGGCACGAAAGGACTTCTCAGTACAC
AGGAACCTCTCCTCACCCAGCGACCTCTCCTCCTGCAGGGAGCTCCCGCCCCCCCATGCAGC
TGGACTCCACCTCAGCAGAGGACACCAGTCCAGCTCTCAGCAGTGGCAGCTCTAAGCCCAGG
GTGTCCATCCCGATGGTCCGCATACTGGCCCCAGTCCTGGTGCTGCTGAGCCTTCTGTCAGC
CGCAGGCCTGATCGCCTTCTGCAGCCACCTGCTCCTGTGGAGAAAGGAAGCTCAACAGGCCA
CGGAGACACAGAGGAACGAGAAGTTCTGGCTCTCACGCTTGACTGCGGAGGAAAAGGAAGCC
CCTTCCCAGGCCCCTGAGGGGGACGTGATCTCGATGCCTCCCCTCCACACATCTGAGGAGGA
GCTGGGCTTCTCGAAGTTTGTCTCAGCGTAGGGCAGGAGGCCCTCCTGGCCAGGCCAGCAGT
GAAGCAGTATGGCTGGCTGGATCAGCACCGATTCCCGAAAGCTTTCCACCTCAGCCTCAGAG
TCCAGCTGCCCGGACTCCAGGGCTCTCCCCACCCTCCCCAGGCTCTCCTCTTGCATGTTCCA
GCCTGACCTAGAAGCGTTTGTCAGCCCTGGAGCCCAGAGCGGTGGCCTTGCTCTTCCGGCTG
GAGACTGGGACATCCCTGATAGGTTCACATCCCTGGGCAGAGTACCAGGCTGCTGACCCTCA
GCAGGGCCAGACAAGGCTCAGTGGATCTGGTCTGAGTTTCAATCTGCCAGGAACTCCTGGGC
CTCATGCCCAGTGTCGGACCCTGCCTTCCTCCCACTCCAGACCCCACCTTGTCTTCCCTCCC
TGGCGTCCTCAGACTTAGTCCCACGGTCTCCTGCATCAGCTGGTGATGAAGAGGAGCATGCT
GGGGTGAGACTGGGATTCTGGCTTCTCTTTGAACCACCTGCATCCAGCCCTTCAGGAAGCCT
GTGAAAAACGTGATTCCTGGCCCCACCAAGACCCACCAAAACCATCTCTGGGCTTGGTGCAG
GACTCTGAATTCTAACAATGCCCAGTGACTGTCGCACTTGAGTTTGAGGGCCAGTGGGCCTG
ATGAACGCTCACACCCCTTCAGCTTAGAGTCTGCATTTGGGCTGTGACGTCTCCACCTGCCC
CAATAGATCTGCTCTGTCTGCGACACCAGATCCACGTGGGGACTCCCCTGAGGCCTGCTAAG
TCCAGGCCTTGGTCAGGTCAGGTGCACATTGCAGGATAAGCCCAGGACCGGCACAGAAGTGG
TTGCCTTTNCCATTTGCCCTCCCTGGNCCATGCCTTCTTGCCTTTGGAAAAAATGATGAAGA
AAACCTTGGCTCCTTCCTTGTCTGGAAAGGGTTACTTGCCTATGGGTTCTGGTGGCTAGAGA
GAAAAGTAGAAAACCAGAGTGCACGTAGGTGTCTAACACAGAGGAGAGTAGGAACAGGGCGG
ATACCTGAAGGTGACTCCGAGTCCAGCCCCCTGGAGAAGGGGTCGGGGGTGGTGGTAAAGTA
GCACAACTACTATTTTTTTTCTTTTTCCATTATTATTGTTTTTTAAGACAGAATCTCGTGCT
GCTGCCCAGGCTGGAGTGCAGTGGCACGATCTGCAAACTCCGCCTCCTGGGTTCAAGTGATT
CTTCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCACGCACCACCACACCTGGCTAATT
TTTGTACTTTTAGTAGAGATGGGGTTTCACCATGTTGGCCAGGCTGGTCTTGAACTCCTGAC
CTCAAATGAGCCTCCTGCTTCAGTCTCCCAAATTGCCGGGATTACAGGCATGAGCCACTGTG
TCTGGCCCTATTTCCTTTAAAAAGTGAAATTAAGAGTTGTTCAGTATGCAAAACTTGGAAAG
ATGGAGGAGAAAAAGAAAAGGAAGAAAAAAATGTCACCCATAGTCTCACCAGAGACTATCAT
TATTTCGTTTTGTTGTACTTCCTTCCACTCTTTTCTTCTTCACATAATTTGCCGGTGTTCTT
TTTACAGAGCAATTATCTTGTATATACAACTTTGTATCCTGCCTTTTCCACCTTATCGTTCC
ATCACTTTATTCCAGCACTTCTCTGTGTTTTACAGACCTTTTTATAAATAAAATGTTCATCA
GCTGCATAAAAAAAAAAAAA

FIGURE 79

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44196
<subunit 1 of 1, 332 aa, 1 stop
<MW: 36143, pI: 5.89, NX(S/T): 1
MRLLVLLWGCLLLPGYEALEGPEEISGFEGDTVSLQCTYREELRDHRKYWCRKGGILFSRCS
GTIYAEEEGQETMKGRVSIRDSRQELSLIVTLWNLTLQDAGEYWCGVEKRGPDESLLISLFV
FPGPCCPPSPSPTFQPLATTRLQPKAKAQQTQPPGLTSPGLYPAATTAKQGKTGAEAPPLPG
TSQYGHERTSQYTGTSPHPATSPPAGSSRPPMQLDSTSAEDTSPALSSGSSKPRVSIPMVRI
LAPVLVLLSLLSAAGLIAFCSHLLLWRKEAQQATETQRNEKFWLSRLTAEEKEAPSQAPEGD
VISMPPLHTSEEELGFSKFVSA

Important features:

Signal peptide:
amino acids 1-17

Transmembrane domain:
amino acids 248-269

N-glycosylation site.
amino acids 96-99

Fibrinogen beta and gamma chains C-terminal domain.
amino acids 104-113

Ig like V-type domain:
amino acids 13-128

FIGURE 80

TTGTGACTAAAAGCTGGCCTAGCAGGCCAGGGAGTGCAGCTGCAGGCGTGGGGGTGGCAGGA
GCCGCAGAGCCAGAGCAGACAGCCGAGAAACAGGTGGACAGTGTGAAAGAACCAGTGGTCTC
GCTCTGTTGCCCAGGCTAGAGTGTACTGGCGTGATCATAGCTCACTGCAGCCTCAGACTCCT
GGACTTGAGAAATCCTCCTGCCTTAGCCTCCTGCATATCTGGGACTCCAGGGGTGCACTCAA
GCCCTGTTTCTTCTCCTTCTGTGAGTGGACCACGGAGGCTGGTGAGCTGCCTGTCATCCCAA
AGCTCAGCTCTGAGCCAGAGTGGTGGTGGCTCCACCTCTGCCGCCGGCATAGAAGCCAGGAG
CAGGGCTCTCAGAAGGCGGTGGTGCCCAGCTGGGATCATGTTGTTGGCCCTGGTCTGTCTGC
TCAGCTGCCTGCTACCCTCCAGTGAGGCCAAGCTCTACGGTCGTTGTGAACTGGCCAGAGTG
CTACATGACTTCGGGCTGGACGGATACCGGGGATACAGCCTGGCTGACTGGGTCTGCCTTGC
TTATTTCACAAGCGGTTTCAACGCAGCTGCTTTGGACTACGAGGCTGATGGGAGCACCAACA
ACGGGATCTTCCAGATCAACAGCCGGAGGTGGTGCAGCAACCTCACCCCGAACGTCCCCAAC
GTGTGCCGGATGTACTGCTCAGATTTGTTGAATCCTAATCTCAAGGATACCGTTATCTGTGC
CATGAAGATAACCCAAGAGCCTCAGGGTCTGGGTTACTGGGAGGCCTGGAGGCATCACTGCC
AGGGAAAAGACCTCACTGAATGGGTGGATGGCTGTGACTTCTAGGATGGACGGAACCATGCA
CAGCAGGCTGGGAAATGTGGTTTGGTTCCTGACCTAGGCTTGGGAAGACAAGCCAGCGAATA
AAGGATGGTTGAACGTGAAA

FIGURE 81

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52187
<subunit 1 of 1, 146 aa, 1 stop
<MW: 16430, pI: 5.05, NX(S/T): 1
MLLALVCLLSCLLPSSEAKLYGRCELARVLHDFGLDGYRGYSLADWVCLAYFTSGFNAAALD
YEADGSTNNGIFQINSRRWCSNLTPNVPNVCRMYCSDLLNPNLKDTVICAMKITQEPQGLGY
WEAWRHHCQGKDLTEWVDGCDF

Important features:

Signal peptide:
amino acids 1-18

N-myristoylation site.
amino acids 67-72

Homolgous region to Alpha-lactalbumin / lysozyme C proteins.
amino acids 34-58 (catalytic domain), 111-132 and 66-107

FIGURE 82

AGCCGCTGCCCCGGGCCGGGCGCCCGCGGCGGCACCATGAGTCCCCGCTCGTGCCTGCGTTC
GCTGCGCCTCCTCGTCTTCGCCGTCTTCTCAGCCGCCGCGAGCAACTGGCTGTACCTGGCCA
AGCTGTCGTCGGTGGGGAGCATCTCAGAGGAGGAGACGTGCGAGAAACTCAAGGGCCTGATC
CAGAGGCAGGTGCAGATGTGCAAGCGGAACCTGGAAGTCATGGACTCGGTGCGCCGCGGTGC
CCAGCTGGCCATTGAGGAGTGCCAGTACCAGTTCCGGAACCGGCGCTGGAACTGCTCCACAC
TCGACTCCTTGCCCGTCTTCGGCAAGGTGGTGACGCAAGGGACTCGGGAGGCGGCCTTCGTG
TACGCCATCTCTTCGGCAGGTGTGGCCTTTGCAGTGACGCGGGCGTGCAGCAGTGGGGAGCT
GGAGAAGTGCGGCTGTGACAGGACAGTGCATGGGGTCAGCCCACAGGGCTTCCAGTGGTCAG
GATGCTCTGACAACATCGCCTACGGTGTGGCCTTCTCACAGTCGTTTGTGGATGTGCGGGAG
AGAAGCAAGGGGGCCTCGTCCAGCAGAGCCCTCATGAACCTCCACAACAATGAGGCCGGCAG
GAAGGCCATCCTGACACACATGCGGGTGGAATGCAAGTGCCACGGGGTGTCAGGCTCCTGTG
AGGTAAAGACGTGCTGGCGAGCCGTGCCGCCCTTCCGCCAGGTGGGTCACGCACTGAAGGAG
AAGTTTGATGGTGCCACTGAGGTGGAGCCACGCCGCGTGGGCTCCTCCAGGGCACTGGTACC
ACGCAACGCACAGTTCAAGCCGCACACAGATGAGGACCTGGTGTACTTGGAGCCTAGCCCCG
ACTTCTGTGAGCAGGACATGCGCAGCGGCGTGCTGGGCACGAGGGGCCGCACATGCAACAAG
ACGTCCAAGGCCATCGACGGCTGTGAGCTGCTGTGCTGTGGCCGCGGCTTCCACACGGCGCA
GGTGGAGCTGGCTGAACGCTGCAGCTGCAAATTCCACTGGTGCTGCTTCGTCAAGTGCCGGC
AGTGCCAGCGGCTCGTGGAGTTGCACACGTGCCGATGACCGCCTGCCTAGCCCTGCGCCGGC
AACCACCTAGTGGCCCAGGGAAGGCCGATAATTTAAACAGTCTCCCACCACCTACCCCAAGA
GATACTGGTTGTATTTTTTGTTCTGGTTTGGTTTTTGGGTCCTCATGTTATTTATTGCCGAA
ACCAGGCAGGCAACCCCAAGGGCACCAACCAGGGCCTCCCCAAAGCCTGGGCCTTTGTGGCT
GCCACTGACCAAAGGGACCTTGCTCGTGCCGCTGGCTGCCCGCATGTGGCTGCCACTGACCA
CTCAGTTGTTATCTGTGTCCGTTTTTCTACTTGCAGACCTAAGGTGGAGTAACAAGGAGTAT
TACCACCACATGGCTACTGACCGTGTCATCGGGGAAGAGGGGGCCTTATGGCAGGGAAAATA
GGTACCGACTTGATGGAAGTCACACCCTCTGGAAAAAAGAACTCTTAACTCTCCAGCACACA
TACACATGGACTCCTGGCAGCTTGAGCCTAGAAGCCATGTCTCTCAAATGCCCTGAGAAAGG
GAACAAGCAGATACCAGGTCAAGGGCACCAGGTTCATTTCAGCCCTTACATGGACAGCTAGA
GGTTCGATATCTGTGGGTCCTTCCAGGCAAGAAGAGGGAGATGAGAGCAAGAGACGACTGAA
GTCCCACCCTAGAACCCAGCCTGCCCCAGCCTGCCCCTGGGAAGAGGAAACTTAACCACTCC
CCAGACCCACCTAGGCAGGCATATAGGCTGCCATCCTGGACCAGGGATCCCGGCTGTGCCTT
TGCAGTCATGCCCGAGTCACCTTTCACAGCGCTGTTCCTCCATGAAACTGAAAAACACACAC
ACACACACACACACACACACACACACACACACACACACACGGACACACACACACACCTGCGAGA
GAGAGGGAGGAAAGGGCTGTGCCTTTGCAGTCATGCCCGAGTCACCTTTCACAGCACTGTTCCTC

FIGURE 83

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48328

<subunit 1 of 1, 351 aa, 1 stop

<MW: 39052, pI: 8.97, NX(S/T): 2

MSPRSCLRSLRLLVFAVFSAAASNWLYLAKLSSVGSISEEETCEKLKGLIQRQVQMCKRNLE
VMDSVRRGAQLAIEECQYQFRNRRWNCSTLDSLPVFGKVVTQGTREAAFVYAISSAGVAFAV
TRACSSGELEKCGCDRTVHGVSPQGFQWSGCSDNIAYGVAFSQSFVDVRERSKGASSSRALM
NLHNNEAGRKAILTHMRVECKCHGVSGSCEVKTCWRAVPPFRQVGHALKEKFDGATEVEPRR
VGSSRALVPRNAQFKPHTDEDLVYLEPSPDFCEQDMRSGVLGTRGRTCNKTSKAIDGCELLC
CGRGFHTAQVELAERCSCKFHWCCFVKCRQCQRLVELHTCR

Important features:

Signal peptide:

amino acids 1-22

N-glycosylation sites.

amino acids 88-91 and 297-300

Wnt-1 family signature.

amino acids 206-215

Homologous region to Wnt-1 family proteins

amino acids 183-235, 305-350, 97-138, 53-92 and 150 -174

FIGURE 84

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCTGGGTGCCTGCAT

CGCC**ATG**GACACCACCAGGTACAGCAAGTGGGGCGGCAGCTCCGAGGAGGTCCCCGGAGGGC

CCTGGGGACGCTGGGTGCACTGGAGCAGGAGACCCCTCTTCTTGGCCCTGGCTGTCCTGGTC

ACCACAGTCCTTTGGGCTGTGATTCTGAGTATCCTATTGTCCAAGGCCTCCACGGAGCGCGC

GGCGCTGCTTGACGGCCACGACCTGCTGAGGACAAACGCCTCGAAGCAGACGGCGGCGCTGG

GTGCCCTGAAGGAGGAGGTCGGAGACTGCCACAGCTGCTGCTCGGGGACGCAGGCGCAGCTG

CAGACCACGCGCGCGGAGCTTGGGGAGGCGCAGGCGAAGCTGATGGAGCAGGAGAGCGCCCT

GCGGGAACTGCGTGAGCGCGTGACCCAGGGCTTGGCTGAAGCCGGCAGGGGCCGTGAGGACG

TCCGCACTGAGCTGTTCCGGGCGCTGGAGGCCGTGAGGCTCCAGAACAACTCCTGCGAGCCG

TGCCCCACGTCGTGGCTGTCCTTCGAGGGCTCCTGCTACTTTTTCTCTGTGCCAAAGACGAC

GTGGGCGGCGGCGCAGGATCACTGCGCAGATGCCAGCGCGCACCTGGTGATCGTTGGGGGCC

TGGATGAGCAGGGCTTCCTCACTCGGAACACGCGTGGCCGTGGTTACTGGCTGGGCCTGAGG

GCTGTGCGCCATCTGGGCAAGGTTCAGGGCTACCAGTGGGTGGACGGAGTCTCTCTCAGCTT

CAGCCACTGGAACCAGGGAGAGCCCAATGACGCTTGGGGGCGCGAGAACTGTGTCATGATGC

TGCACACGGGGCTGTGGAACGACGCACCGTGTGACAGCGAGAAGGACGGCTGGATCTGTGAG

AAAAGGCACAACTGC**TGA**CCCCGCCCAGTGCCCTGGAGCCGCGCCCATTGCAGCATGTCGTA

TCCTGGGGGCTGCTCACCTCCCTGGCTCCTGGAGCTGATTGCCAAAGAGTTTTTTTCTTCCT

CATCCACCGCTGCTGAGTCTCAGAAACACTTGGCCCAACATAGCCCTGTCCAGCCCAGTGCC

TGGGCTCTGGGACCTCCATGCCGACCTCATCCTAACTCCACTCACGCAGACCCAACCTAACC

TCCACTAGCTCCAAAATCCCTGCTCCTGCGTCCCCGTGATATGCCTCCACTTCTCTCCCTAA

CCAAGGTTAGGTGACTGAGGACTGGAGCTGTTTGGTTTTCTCGCATTTTCCACCAAACTGGA

AGCTGTTTTTGCAGCCTGAGGAAGCATCAATAAATATTTGAGAAATGAAAAAA

FIGURE 85

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56352

<subunit 1 of 1, 293 aa, 1 stop

<MW: 32562, pI: 6.53, NX(S/T): 2

MDTTRYSKWGGSSEEVPGGPWGRWVHWSRRPLFLALAVLVTTVLWAVILSILLSKASTERAA
LLDGHDLLRTNASKQTAALGALKEEVGDCHSCCSGTQAQLQTTRAELGEAQAKLMEQESALR
ELRERVTQGLAEAGRGREDVRTELFRALEAVRLQNNSCEPCPTSWLSFEGSCYFFSVPKTTW
AAAQDHCADASAHLVIVGGLDEQGFLTRNTRGRGYWLGLRAVRHLGKVQGYQWVDGVSLSFS
HWNQGEPNDAWGRENCVMMLHTGLWNDAPCDSEKDGWICEKRHNC

Important features:

Type II transmembrane domain:
amino acids 31-54

N-glycosylation sites.
amino acids 73-76 and 159-162

Leucine zipper pattern.
amino acids 102-123

N-myristoylation sites.
amino acids 18-23, 133-138 and 242-247

C-type lectin domain signature.
amino acids 264-287

FIGURE 86

GCCAGGGGAAGAGGGTGATCCGACCCGGGGAAGGTCGCTGGGCAGGGCGAGTTGGGAAAGCG
GCAGCCCCCGCCGCCCCCGCAGCCCCTTCTCCTCCTTTCTCCCACGTCCTATCTGCCTCTCG
CTGGAGGCCAGGCCGTGCAGCATCGAAGACAGGAGGAACTGGAGCCTCATTGGCCGGCCCGG
GGCGCCGGCCTCGGGCTTAAATAGGAGCTCCGGGCTCTGGCTGGGACCCGACCGCTGCCGGC
CGCGCTCCCGCTGCTCCTGCCGGGTGATGGAAAACCCCAGCCCGGCCGCCGCCCTGGGCAAG
GCCCTCTGCGCTCTCCTCCTGGCCACTCTCGGCGCCGCCGGCCAGCCTCTTGGGGGAGAGTC
CATCTGTTCCGCCAGAGCCCCGGCCAAATACAGCATCACCTTCACGGGCAAGTGGAGCCAGA
CGGCCTTCCCCAAGCAGTACCCCCTGTTCCGCCCCCCTGCGCAGTGGTCTTCGCTGCTGGGG
GCCGCGCATAGCTCCGACTACAGCATGTGGAGGAAGAACCAGTACGTCAGTAACGGGCTGCG
CGACTTTGCGGAGCGCGGCGAGGCCTGGGCGCTGATGAAGGAGATCGAGGCGGCGGGGGAGG
CGCTGCAGAGCGTGCACGAGGTGTTTTCGGCGCCCGCCGTCCCCAGCGGCACCGGGCAGACG
TCGGCGGAGCTGGAGGTGCAGCGCAGGCACTCGCTGGTCTCGTTTGTGGTGCGCATCGTGCC
CAGCCCCGACTGGTTCGTGGGCGTGGACAGCCTGGACCTGTGCGACGGGGACCGTTGGCGGG
AACAGGCGGCGCTGGACCTGTACCCCTACGACGCCGGGACGGACAGCGGCTTCACCTTCTCC
TCCCCCAACTTCGCCACCATCCCGCAGGACACGGTGACCGAGATAACGTCCTCCTCTCCCAG
CCACCCGGCCAACTCCTTCTACTACCCGCGGCTGAAGGCCCTGCCTCCCATCGCCAGGGTGA
CACTGCTGCGGCTGCGACAGAGCCCCAGGGCCTTCATCCCTCCCGCCCCAGTCCTGCCCAGC
AGGGACAATGAGATTGTAGACAGCGCCTCAGTTCCAGAAACGCCGCTGGACTGCGAGGTCTC
CCTGTGGTCGTCCTGGGGACTGTGCGGAGGCCACTGTGGGAGGCTCGGGACCAAGAGCAGGA
CTCGCTACGTCCGGGTCCAGCCCGCCAACAACGGGAGCCCCTGCCCCGAGCTCGAAGAAGAG
GCTGAGTGCGTCCCTGATAACTGCGTCTAAGACCAGAGCCCCGCAGCCCCTGGGGCCCCCCG
GAGCCATGGGGTGTCGGGGGCTCCTGTGCAGGCTCATGCTGCAGGCGGCCGAGGGCACAGGG
GGTTTCGCGCTGCTCCTGACCGCGGTGAGGCCGCGCCGACCATCTCTGCACTGAAGGCCCT
CTGGTGGCCGGCACGGGCATTGGGAAACAGCCTCCTCCTTTCCCAACCTTGCTTCTTAGGGG
CCCCCGTGTCCCGTCTGCTCTCAGCCTCCTCCTCCTGCAGGATAAAGTCATCCCCAAGGCTC
CAGCTACTCTAAATTATGTCTCCTTATAAGTTATTGCTGCTCCAGGAGATTGTCCTTCATCG
TCCAGGGGCCTGGCTCCCACGTGGTTGCAGATACCTCAGACCTGGTGCTCTAGGCTGTGCTG
AGCCCACTCTCCCGAGGGCGCATCCAAGCGGGGGCCACTTGAGAAGTGAATAAATGGGGCGG
TTTCGGAAGCGTCAGTGTTTCCATGTTATGGATCTCTCTGCGTTTGAATAAAGACTATCTCT
GTTGCTCACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 87

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53971

><subunit 1 of 1, 331 aa, 1 stop

><MW: 35844, pI: 5.45, NX(S/T): 2

MENPSPAAALGKALCALLLATLGAAGQPLGGESICSARAPAKYSITFTGKWSQTAFPKQYPL
FRPPAQWSSLLGAAHSSDYSMWRKNQYVSNGLRDFAERGEAWALMKEIEAAGEALQSVHEVF
SAPAVPSGTGQTSAELEVQRRHSLVSFVVRIVPSPDWFVGVDSLDLCDGDRWREQAALDLYP
YDAGTDSGFTFSSPNFATIPQDTVTEITSSSPSHPANSFYYPRLKALPPIARVTLLRLRQSP
RAFIPPAPVLPSRDNEIVDSASVPETPLDCEVSLWSSWGLCGGHCGRLGTKSRTRYVRVQPA
NNGSPCPELEEEAECVPDNCV

Important features:

Signal peptide:

amino acids 1-26

FIGURE 88

GGCGGCGTCCGTGAGGGGCTCCTTTGGGCAGGGGTAGTGTTTGGTGTCCCTGTCTTGCGTGA
TATTGACAAACTGAAGCTTTCCTGCACCACTGGACTTAAGGAAGAGTGTACTCGTAGGCGGA
CAGCTTTAGTGGCCGGCCGGCCGCTCTCATCCCCCGTAAGGAGCAGAGTCCTTTGTACTGAC
CAAGATGAGCAACATCTACATCCAGGAGCCTCCCACGAATGGGAAGGTTTTATTGAAAACTA
CAGCTGGAGATATTGACATAGAGTTGTGGTCCAAAGAAGCTCCTAAAGCTTGCAGAAATTTT
ATCCAACTTTGTTTGGAAGCTTATTATGACAATACCATTTTTCATAGAGTTGTGCCTGGTTT
CATAGTCCAAGGCGGAGATCCTACTGGCACAGGGAGTGGTGGAGAGTCTATCTATGGAGCGC
CATTCAAAGATGAATTTCATTCACGGTTGCGTTTTAATCGGAGAGGACTGGTTGCCATGGCA
AATGCTGGTTCTCATGATAATGGCAGCCAGTTTTTCTTCACACTGGGTCGAGCAGATGAACT
TAACAATAAGCATACCATCTTTGGAAAGGTTACAGGGGATACAGTATATAACATGTTGCGAC
TGTCAGAAGTAGACATTGATGATGACGAAAGACCACATAATCCACACAAAATAAAAAGCTGT
GAGGTTTTGTTTAATCCTTTTGATGACATCATTCCAAGGGAAATTAAAAGGCTGAAAAAAGA
GAAACCAGAGGAGGAAGTAAAGAAATTGAAACCCAAAGGCACAAAAAATTTTAGTTTACTTT
CATTTGGAGAGGAAGCTGAGGAAGAAGAGGAGGAAGTAAATCGAGTTAGTCAGAGCATGAAG
GGCAAAAGCAAAAGTAGTCATGACTTGCTTAAGGATGATCCACATCTCAGTTCTGTTCCAGT
TGTAGAAAGTGAAAAAGGTGATGCACCAGATTTAGTTGATGATGGAGAAGATGAAAGTGCAG
AGCATGATGAATATATTGATGGTGATGAAAAGAACCTGATGAGAGAAAGAATTGCCAAAAAA
TTAAAAAAGGACACAAGTGCGAATGTTAAATCAGCTGGAGAAGGAGAAGTGGAGAAGAAATC
AGTCAGCCGCAGTGAAGAGCTCAGAAAAGAAGCAAGACAATTAAAACGGGAACTCTTAGCAG
CAAAACAAAAAAAAAGTAGAAAATGCAGCAAAACAAGCAGAAAAAAGAAGTGAAGAGGAAGAA
GCCCCTCCAGATGGTGCTGTTGCCGAATACAGAAGAGAAAAGCAAAAGTATGAAGCTTTGAG
GAAGCAACAGTCAAAGAAGGGAACTTCCCGGGAAGATCAGACCCTTGCACTGCTGAACCAGT
TTAAATCTAAACTCACTCAAGCAATTGCTGAAACACCTGAAAATGACATTCCTGAAACAGAA
GTAGAAGATGATGAAGGATGGATGTCACATGTACTTCAGTTTGAGGATAAAAGCAGAAAAGT
GAAAGATGCAAGCATGCAAGACTCAGATACATTTGAAATCTATGATCCTCGGAATCCAGTGA
ATAAAAGAAGGAGGGAAGAAAGCAAAAAGCTGATGAGAGAGAAAAAAGAAAGAAGATAAAAT
GAGAATAATGATAACCAGAACTTGCTGGAAATGTGCCTACAATGGCCTTGTAACAGCCATTG
TTCCCAACAGCATCACTTAGGGGTGTGAAAAGAAGTATTTTTGAACCTGTTGTCTGGTTTTG
AAAAACAATTATCTTGTTTTGCAAATTGTGGAATGATGTAAGCAAATGCTTTTGGTTACTGG
TACATGTGTTTTTTCCTAGCTGACCTTTTATATTGCTAAATCTGAAATAAAATAACTTTCCT
TCCACAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 89

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50919
><subunit 1 of 1, 472 aa, 1 stop
><MW: 53847, pI: 5.75, NX(S/T): 2
MSNIYIQEPPTNGKVLLKTTAGDIDIELWSKEAPKACRNFIQLCLEAYYDNTIFHRVVPGFI
VQGGDPTGTGSGGESIYGAPFKDEFHSRLRFNRRGLVAMANAGSHDNGSQFFFTLGRADELN
NKHTIFGKVTGDTVYNMLRLSEVDIDDDERPHNPHKIKSCEVLFNPFDDIIPREIKRLKKEK
PEEEVKKLKPKGTKNFSLLSFGEEAEEEEEEVNRVSQSMKGKSKSSHDLLKDDPHLSSVPVV
ESEKGDAPDLVDDGEDESAEHDEYIDGDEKNLMRERIAKKLKKDTSANVKSAGEGEVEKKSV
SRSEELRKEARQLKRELLAAKQKKVENAAKQAEKRSEEEEAPPDGAVAEYRREKQKYEALRK
QQSKKGTSREDQTLALLNQFKSKLTQAIAETPENDIPETEVEDDEGWMSHVLQFEDKSRKVK
DASMQDSDTFEIYDPRNPVNKRRREESKKLMREKKERR

Important features:

Signal peptide:
amino acids 1-21

N-glycosylation sites.
amino acids 109-112 and 201-204

Cyclophilin-type peptidyl-prolyl cis-trans isomerase signature.
amino acids 49-66

Homologous region to Cyclophilin-type peptidyl-prolyl cis-trans isomerase
amino acids 96-140, 49-89 and 22-51

FIGURE 90

CGCCGCCGTTGGGGCTGGAAGTTCCCGCCAGGTCCGTGCCGGGCGAGAGAGATGCTGCCCGG
CCCGCCTCGGCTTTGAGGCGAGAGAAGTGTCCCAGACCCATTTCGCCTTGCTGACGGCGTCG
AGCCCTGGCCAGACATGTCCACAGGGTTCTCCTTCGGGTCCGGGACTCTGGGCTCCACCACC
GTGGCCGCCGGCGGGACCAGCACAGGCGGCGTTTTCTCCTTCGGAACGGGAACGTCTAGCAA
CCCTTCTGTGGGGCTCAATTTTGGAAATCTTGGAAGTACTTCAACTCCAGCAACTACATCTG
CTCCTTCAAGTGGTTTTGGAACCGGGCTCTTTGGATCTAAACCTGCCACTGGGTTCACTCTA
GGAGGAACAAATACAGGTGCCTTGCACACCAAGAGGCCTCAAGTGGTCACCAAATATGGAAC
CCTGCAAGGAAAACAGATGCATGTGGGGAAGACACCCATCCAAGTCTTTTTAGGAGTCCCCT
TCTCCAGACCTCCTCTAGGTATCCTCAGGTTTGCACCTCCAGAACCCCCGGAGCCCTGGAAA
GGAATCAGAGATGCTACCACCTACCCGCCTGGATGGAGTCTCGCTCTGTCGCCAGGCTGGAG
TGCAGTGGCACGATCTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGCGAGTCTCCTGC
CTCAGCCTCTGAGTGTCTGGGGCTACAGGTGCCTGCAGGAGTCCTGGGGCCAGCTGGCCTCG
ATGTACGTCAGCACGCGGGAACGGTACAAGTGGCTGCGCTTCAGCGAGGACTGTCTGTACCT
GAACGTGTACGCGCCGGCGCGCGCGCCCGGGGATCCCCAGCTGCCAGTGATGGTCTGGTTCC
CGGGAGGCGCCTTCATCGTGGGCGCTGCTTCTTCGTACGAGGGCTCTGACTTGGCCGCCCGC
GAGAAAGTGGTGCTGGTGTTTCTGCAGCACAGGCTCGGCATCTTCGGCTTCCTGAGCACGGA
CGACAGCCACGCGCGCGGGAACTGGGGGCTGCTGGACCAGATGGCGGCTCTGCGCTGGGTGC
AGGAGAACATCGCAGCCTTCGGGGGAGACCCAGGAAATGTGACCCTGTTCGGCCAGTCGGCG
GGGGCCATGAGCATCTCAGGACTGATGATGTCACCCCTAGCCTCGGGTCTCTTCCATCGGGC
CATTTCCCAGAGTGGCACCGCGTTATTCAGACTTTTCATCACTAGTAACCCACTGAAAGTGG
CCAAGAAGGTTGCCCACCTGGCTGGATGCAACCACAACAGCACACAGATCCTGGTAAACTGC
CTGAGGGCACTATCAGGGACCAAGGTGATGCGTGTGTCCAACAAGATGAGATTCCTCCAACT
GAACTTCCAGAGAGACCCGGAAGAGATTATCTGGTCCATGAGCCCTGTGGTGGATGGTGTGG
TGATCCCAGATGACCCTTTGGTGCTCCTGACCCAGGGGAAGGTTTCATCTGTGCCCTACCTT
CTAGGTGTCAACAACCTGGAATTCAATTGGCTCTTGCCTTATAATATCACCAAGGAGCAGGT
ACCACTTGTGGTGGAGGAGTACCTGGACAATGTCAATGAGCATGACTGGAAGATGCTACGAA
ACCGTATGATGGACATAGTTCAAGATGCCACTTTCGTGTATGCCACACTGCAGACTGCTCAC
TACCACCGAGAAACCCCAATGATGGGAATCTGCCCTGCTGGCCACGCTACAACAAGGATGAA
AAGTACCTGCAGCTGGATTTTACCACAAGAGTGGGCATGAAGCTCAAGGAGAAGAAGATGGC
TTTTTGGATGAGTCTGTACCAGTCTCAAAGACCTGAGAAGCAGAGGCAATTCTAAGGGTGGC
TATGCAGGAAGGAGCCAAAGAGGGGTTTGCCCCCACCATCCAGGCCCTGGGGAGACTAGCCA
TGGACATACCTGGGGACAAGAGTTCTACCCACCCCAGTTTAGAACTGCAGGAGCTCCCTGCT
GCCTCCAGGCCAAAGCTAGAGCTTTTGCCTGTTGTGTGGGACCTGCACTGCCCTTTCCAGCC
TGACATCCCATGATGCCCCTCTACTTCACTGTTGACATCCAGTTAGGCCAGGCCCTGTCAAC
ACCACACTGTGCTCAGCTCTCCAGCCTCAGGACAACCTCTTTTTTTCCCTTCTTCAAATCCT
CCCACCCTTCAATGTCTCCTTGTGACTCCTTCTTATGGGAGGTCGACCCAGACTGCCACTGC
CCCTGTCACTGCACCCAGCTTGGCATTTACCATCCATCCTGCTCAACCTTGTTCCTGTCTGT
TCACATTGGCCTGGAGGCCTAGGGCAGGTTGTGACATGGAGCAAACTTTTGGTAGTTTGGGA
TCTTCTCTCCCACCCACACTTATCTCCCCCAGGGCCACTCCAAAGTCTATACACAGGGGTGG
TCTCTTCAATAAAGAAGTGTTGATTAGAAAAAAAAAAA

FIGURE 91

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44179

<subunit 1 of 1, 545 aa, 1 stop

<MW: 58934, pI: 9.45, NX(S/T): 4

MSTGFSFGSGTLGSTTVAAGGTSTGGVFSFGTGTSSNPSVGLNFGNLGSTSTPATTSAPSSG
FGTGLFGSKPATGFTLGGTNTGALHTKRPQVVTKYGTLQGKQMHVGKTPIQVFLGVPFSRPP
LGILRFAPPEPPEPWKGIRDATTYPPGWSLALSPGWSAVARSRLTATSASRVQASLLPQPLS
VWGYRCLQESWGQLASMYVSTRERYKWLRFSEDCLYLNVYAPARAPGDPQLPVMVWFPGGAF
IVGAASSYEGSDLAAREKVVLVFLQHRLGIFGFLSTDDSHARGNWGLLDQMAALRWVQENIA
AFGGDPGNVTLFGQSAGAMSISGLMMSPLASGLFHRAISQSGTALFRLFITSNPLKVAKKVA
HLAGCNHNSTQILVNCLRALSGTKVMRVSNKMRFLQLNFQRDPEEIIWSMSPVVDGVVIPDD
PLVLLTQGKVSSVPYLLGVNNLEFNWLLPYNITKEQVPLVVEEYLDNVNEHDWKMLRNRMMD
IVQDATFVYATLQTAHYHRETPMMGICPAGHATTRMKSTCSWILPQEWA

Important features:

Signal peptide:

amino acids 1-29

Carboxylesterases type-B serine active site.

amino acids 312-327

Carboxylesterases type-B signature 2.

amino acids 218-228

N-glycosylation sites.

amino acids 318-321, 380-383 and 465-468

FIGURE 92

GAGAACAGGCCTGTCTCAGGCAGGCCCTGCGCCTCCTATGCGGAGATGCTACTGCCACTGCT
GCTGTCCTCGCTGCTGGGCGGGTCCCAGGCTATGGATGGGAGATTCTGGATACGAGTGCAGG
AGTCAGTGATGGTGCCGGAGGGCCTGTGCATCTCTGTGCCCTGCTCTTTCTCCTACCCCCGA
CAAGACTGGACAGGGTCTACCCCAGCTTATGGCTACTGGTTCAAAGCAGTGACTGAGACAAC
CAAGGGTGCTCCTGTGGCCACAAACCACCAGAGTCGAGAGGTGGAAATGAGCACCCGGGGCC
GATTCCAGCTCACTGGGGATCCCGCCAAGGGGAACTGCTCCTTGGTGATCAGAGACGCGCAG
ATGCAGGATGAGTCACAGTACTTCTTTCGGGTGGAGAGAGGAAGCTATGTGACATATAATTT
CATGAACGATGGGTTCTTTCTAAAAGTAACAGTGCTCAGCTTCACGCCCAGACCCCAGGACC
ACAACACCGACCTCACCTGCCATGTGGACTTCTCCAGAAAGGGTGTGAGCGCACAGAGGACC
GTCCGACTCCGTGTGGCCTATGCCCCCAGAGACCTTGTTATCAGCATTTCACGTGACAACAC
GCCAGCCCTGGAGCCCCAGCCCCAGGGAAATGTCCCATACCTGGAAGCCCAAAAAGGCCAGT
TCCTGCGGCTCCTCTGTGCTGCTGACAGCCAGCCCCCTGCCACACTGAGCTGGGTCCTGCAG
AACAGAGTCCTCTCCTCGTCCCATCCCTGGGGCCCTAGACCCCTGGGGCTGGAGCTGCCCGG
GGTGAAGGCTGGGGATTCAGGGCGCTACACCTGCCGAGCGGAGAACAGGCTTGGCTCCCAGC
AGCGAGCCCTGGACCTCTCTGTGCAGTATCCTCCAGAGAACCTGAGAGTGATGGTTTCCCAA
GCAAACAGGACAGTCCTGGAAAACCTTGGGAACGGCACGTCTCTCCCAGTACTGGAGGGCCA
AAGCCTGTGCCTGGTCTGTGTCACACACAGCAGCCCCCCAGCCAGGCTGAGCTGGACCCAGA
GGGGACAGGTTCTGAGCCCCTCCCAGCCCTCAGACCCCGGGGTCCTGGAGCTGCCTCGGGTT
CAAGTGGAGCACGAAGGAGAGTTCACCTGCCACGCTCGGCACCCACTGGGCTCCCAGCACGT
CTCTCTCAGCCTCTCCGTGCACTATAAGAAGGGACTCATCTCAACGGCATTCTCCAACGGAG
CGTTTCTGGGAATCGGCATCACGGCTCTTCTTTTCCTCTGCCTGGCCCTGATCATCATGAAG
ATTCTACCGAAGAGACGGACTCAGACAGAAACCCCGAGGCCCAGGTTCTCCCGGCACAGCAC
GATCCTGGATTACATCAATGTGGTCCCGACGGCTGGCCCCCTGGCTCAGAAGCGGAATCAGA
AAGCCACACCAAACAGTCCTCGGACCCCTCCTCCACCAGGTGCTCCCTCCCCAGAATCAAAG
AAGAACCAGAAAAAGCAGTATCAGTTGCCCAGTTTCCCAGAACCCAAATCATCCACTCAAGC
CCCAGAATCCCAGGAGAGCCAAGAGGAGCTCCATTATGCCACGCTCAACTTCCCAGGCGTCA
GACCCAGGCCTGAGGCCCGGATGCCCAAGGGCACCCAGGCGGATTATGCAGAAGTCAAGTTC
CAATGAGGGTCTCTTAGGCTTTAGGACTGGGACTTCGGCTAGGGAGGAAGGTAGAGTAAGAG
GTTGAAGATAACAGAGTGCAAAGTTTCCTTCTCTCCCTCTCTCTCTCTCTTTCTCTCTCTCT
CTCTCTTTCTCTCTCTTTTAAAAAAACATCTGGCCAGGGCACAGTGGCTCACGCCTGTAATC
CCAGCACTTTGGGAGGTTGAGGTGGGCAGATCGCCTGAGGTCGGGAGTTCGAGACCAGCCTG
GCCAACTTGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGCTGGGCATGGTGGCAGG
CGCCTGTAATCCTACCTACTTGGGAAGCTGAGGCAGGAGAATCACTTGAACCTGGGAGACGG
AGGTTGCAGTGAGCCAAGATCACACCATTGCACGCCAGCCTGGGCAACAAAGCGAGACTCCA
TCTCAAAAAAAAAAATCCTCCAAATGGGTTGGGTGTCTGTAATCCCAGCACTTTGGGAGGCTA
AGGTGGGTGGATTGCTTGAGCCCAGGAGTTCGAGACCAGCCTGGGCAACATGGTGAAACCCC
ATCTCTACAAAAAATACAAAACATAGCTGGGCTTGGTGGTGTGTGCCTGTAGTCCCAGCTGT
CAGACATTTAAACCAGAGCAACTCCATCTGGAATAGGAGCTGAATAAAATGAGGCTGAGACC
TACTGGGCTGCATTCTCAGACAGTGGAGGCATTCTAAGTCACAGGATGAGACAGGAGGTCCG
TACAAGATACAGGTCATAAAGACTTTGCTGATAAAACAGATTGCAGTAAAGAAGCCAACCAA
ATCCCACCAAAACCAAGTTGGCCACGAGAGTGACCTCTGGTCGTCCTCACTGCTACACTCCT
GACAGCACCATGACAGTTTACAAATGCCATGGCAACATCAGGAAGTTACCCGATATGTCCCA
AAAGGGGGAGGAATGAATAATCCACCCCTTGTTTAGCAAATAAGCAAGAAATAACCATAAAA
GTGGGCAACCAGCAGCTCTAGGCGCTGCTCTTGTCTATGGAGTAGCCATTCTTTTGTTCCTT
TACTTTCTTAATAAACTTGCTTTCACCTTAAAAAAA

FIGURE 93

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA54002
><subunit 1 of 1, 544 aa, 1 stop
><MW: 60268, pI: 9.53, NX(S/T): 3
MLLPLLLSSLLGGSQAMDGRFWIRVQESVMVPEGLCISVPCSFSYPRQDWTGSTPAYGYWFK
AVTETTKGAPVATNHQSREVEMSTRGRFQLTGDPAKGNCSLVIRDAQMQDESQYFFRVERGS
YVTYNFMNDGFFLKVTVLSFTPRPQDHNTDLTCHVDFSRKGVSAQRTVRLRVAYAPRDLVIS
ISRDNTPALEPQPQGNVPYLEAQKGQFLRLLCAADSQPPATLSWVLQNRVLSSSHPWGPRPL
GLELPGVKAGDSGRYTCRAENRLGSQQRALDLSVQYPPENLRVMVSQANRTVLENLGNGTSL
PVLEGQSLCLVCVTHSSPPARLSWTQRGQVLSPSQPSDPGVLELPRVQVEHEGEFTCHARHP
LGSQHVSLSLSVHYKKGLISTAFSNGAFLGIGITALLFLCLALIIMKILPKRRTQTETPRPR
FSRHSTILDYINVVPTAGPLAQKRNQKATPNSPRTPPPPGAPSPESKKNQKKQYQLPSFPEP
KSSTQAPESQESQEELHYATLNFPGVRPRPEARMPKGTQADYAEVKFQ

Important features:

Signal peptide:
amino acids 1-15

Transmembrane domain:
amino acids 399-418

N-glycosylation site.
amino acids 100-103, 297-300 and 306-309

Immunoglobulins and major histocompatibility complex proteins signature.
amino acids 365-371

FIGURE 94

TGAAGAGTAATAGTTGGAATCAAAAGAGTCAACGCAATGAACTGTTATTTACTGCTGCGTTT
TATGTTGGGAATTCCTCTCCTATGGCCTTGTCTTGGAGCAACAGAAAACTCTCAAACAAAGA
AAGTCAAGCAGCCAGTGCGATCTCATTTGAGAGTGAAGCGTGGCTGGGTGTGGAACCAATTT
TTTGTACCAGAGGAAATGAATACGACTAGTCATCACATCGGCCAGCTAAGATCTGATTTAGA
CAATGGAAACAATTCTTTCCAGTACAAGCTTTTGGGAGCTGGAGCTGGAAGTACTTTTATCA
TTGATGAAAGAACAGGTGACATATATGCCATACAGAAGCTTGATAGAGAGGAGCGATCCCTC
TACATCTTAAGAGCCCAGGTAATAGACATCGCTACTGGAAGGGCTGTGGAACCTGAGTCTGA
GTTTGTCATCAAAGTTTCGGATATCAATGACAATGAACCAAAATTCCTAGATGAACCTTATG
AGGCCATTGTACCAGAGATGTCTCCAGAAGGAACATTAGTTATCCAGGTGACAGCAAGTGAT
GCTGACGATCCCTCAAGTGGTAATAATGCTCGTCTCCTCTACAGCTTACTTCAAGGCCAGCC
ATATTTTTCTGTTGAACCAACAACAGGAGTCATAAGAATATCTTCTAAAATGGATAGAGAAC
TGCAAGATGAGTATTGGGTAATCATTCAAGCCAAGGACATGATTGGTCAGCCAGGAGCGTTG
TCTGGAACAACAAGTGTATTAATTAAACTTTCAGATGTTAATGACAATAAGCCTATATTTAA
AGAAAGTTTATACCGCTTGACTGTCTCTGAATCTGCACCCACTGGGACTTCTATAGGAACAA
TCATGGCATATGATAATGACATAGGAGAGAATGCAGAAATGGATTACAGCATTGAAGAGGAT
GATTCGCAAACATTTGACATTATTACTAATCATGAAACTCAAGAAGGAATAGTTATATTAAA
AAAGAAAGTGGATTTTGAGCACCAGAACCACTACGGTATTAGAGCAAAAGTTAAAAACCATC
ATGTTCCTGAGCAGCTCATGAAGTACCACACTGAGGCTTCCACCACTTTCATTAAGATCCAG
GTGGAAGATGTTGATGAGCCTCCTCTTTTCCTCCTTCCATATTATGTATTTGAAGTTTTTGA
AGAAACCCCACAGGGATCATTTGTAGGCGTGGTGTCTGCCACAGACCCAGACAATAGGAAAT
CTCCTATCAGGTATTCTATTACTAGGAGCAAAGTGTTCAATATCAATGATAATGGTACAATC
ACTACAAGTAACTCACTGGATCGTGAAATCAGTGCTTGGTACAACCTAAGTATTACAGCCAC
AGAAAAATACAATATAGAACAGATCTCTTCGATCCCACTGTATGTGCAAGTTCTTAACATCA
ATGATCATGCTCCTGAGTTCTCTCAATACTATGAGACTTATGTTTGTGAAAATGCAGGCTCT
GGTCAGGTAATTCAGACTATCAGTGCAGTGGATAGAGATGAATCCATAGAAGAGCACCATTT
TTACTTTAATCTATCTGTAGAAGACACTAACAATTCAAGTTTTACAATCATAGATAATCAAG
ATAACACAGCTGTCATTTTGACTAATAGAACTGGTTTTAACCTTCAAGAAGAACCTGTCTTC
TACATCTCCATCTTAATTGCCGACAATGGAATCCCGTCACTTACAAGTACAAACACCCTTAC
CATCCATGTCTGTGACTGTGGTGACAGTGGGAGCACACAGACCTGCCAGTACCAGGAGCTTG
TGCTTTCCATGGGATTCAAGACAGAAGTTATCATTGCTATTCTCATTTGCATTATGATCATA
TTTGGGTTTATTTTTTTGACTTTGGGTTTAAAACAACGGAGAAAACAGATTCTATTTCCTGA
GAAAAGTGAAGATTTCAGAGAGAATATATTCCAATATGATGATGAAGGGGGTGGAGAAGAAG
ATACAGAGGCCTTTGATATAGCAGAGCTGAGGAGTAGTACCATAATGCGGGAACGCAAGACT
CGGAAAACCACAAGCGCTGAGATCAGGAGCCTATACAGGCAGTCTTTGCAAGTTGGCCCCGA
CAGTGCCATATTCAGGAAATTCATTCTGGAAAAGCTCGAAGAAGCTAATACTGATCCGTGTG
CCCCTCCTTTTGATTCCCTCCAGACCTACGCTTTTGAGGGAACAGGGTCATTAGCTGGATCC
CTGAGCTCCTTAGAATCAGCAGTCTCTGATCAGGATGAAAGCTATGATTACCTTAATGAGTT
GGGACCTCGCTTTAAAAGATTAGCATGCATGTTTGGTTCTGCAGTGCAGTCAAATAATTAGG
GCTTTTTACCATCAAAATTTTTAAAAGTGCTAATGTGTATTCGAACCCAATGGTAGTCTTAA
AGAGTTTTGTGCCCTGGCTCTATGGCGGGGAAAGCCCTAGTCTATGGAGTTTTCTGATTTCC
CTGGAGTAAATACTCCATGGTTATTTTAAGCTACCTACATGCTGTCATTGAACAGAGATGTG
GGGAGAAATGTAAACAATCAGCTCACAGGCATCAATACAACCAGATTTGAAGTAAAATAATG
TAGGAAGATATTAAAAGTAGATGAGAGGACACAAGATGTAGTCGATCCTTATGCGATTATAT
CATTATTTACTTAGGAAAGAGTAAAAATACCAAACGAGAAAATTTAAAGGAGCAAAAATTTG
CAAGTCAAATAGAAATGTACAAATCGAGATAACATTTACATTTCTATCATATTGACATGAAA
ATTGAAAATGTATAGTCAGAGAAATTTTCATGAATTATTCCATGAAGTATTGTTTCCTTTAT
TTAAA

FIGURE 95

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53906

><subunit 1 of 1, 772 aa, 1 stop

><MW: 87002, pI: 4.64, NX(S/T): 8

MNCYLLLRFMLGIPLLWPCLGATENSQTKKVKQPVRSHLRVKRGWVWNQFFVPEEMNTTSHH
IGQLRSDLDNGNNSFQYKLLGAGAGSTFIIDERTGDIYAIQKLDREERSLYILRAQVIDIAT
GRAVEPESEFVIKVSDINDNEPKFLDEPYEAIVPEMSPEGTLVIQVTASDADDPSSGNNARL
LYSLLQGQPYFSVEPTTGVIRISSKMDRELQDEYWVIIQAKDMIGQPGALSGTTSVLIKLSD
VNDNKPIFKESLYRLTVSESAPTGTSIGTIMAYDNDIGENAEMDYSIEEDDSQTFDIITNHE
TQEGIVILKKKVDFEHQNHYGIRAKVKNHHVPEQLMKYHTEASTTFIKIQVEDVDEPPLFLL
PYYVFEVFEETPQGSFVGVVSATDPDNRKSPIRYSITRSKVFNINDNGTITTSNSLDREISA
WYNLSITATEKYNIEQISSIPLYVQVLNINDHAPEFSQYYETYVCENAGSGQVIQTISAVDR
DESIEEHHFYFNLSVEDTNNSSFTIIDNQDNTAVILTNRTGFNLQEEPVFYISILIADNGIP
SLTSTNTLTIHVCDCGDSGSTQTCQYQELVLSMGFKTEVIIAILICIMIIFGFIFLTLGLKQ
RRKQILFPEKSEDFRENIFQYDDEGGGEEDTEAFDIAELRSSTIMRERKTRKTTSAEIRSLY
RQSLQVGPDSAIFRKFILEKLEEANTDPCAPPFDSLQTYAFEGTGSLAGSLSSLESAVSDQD
ESYDYLNELGPRFKRLACMFGSAVQSNN

Important features:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 597-617

N-glycosylation sites.

amino acids 57-60, 74-77, 419-423, 437-440, 508-511, 515-518, 516-519 and 534-537

Cadherins extracellular repeated domain signature.

amino acids 136-146 and 244-254

FIGURE 96

ATTTCAAGGCCAGCCATATTTTTNTGTTGAACCAACAACAGGAGTCATAAGAATATTTTNTA
AAATGGATAGAGAACTGCAAGATGAGTATTGGGTAATCATTCAAGCCAAGGACATGATTGGT
CAGCCAGGAGCGTTGTNTGGAACAACAAGTGTATTAATTAAACTTTCAGATGTTAATGACAA
TAAGCCTATATTTAAAGAAAGTTTATACCGCTTGACTGTNTNTGAATCTGCACCCACTGGGA
NTTNTATAGGAACAATCATGGCATATGATAATGACATAGGAGAGAATGCAGAAATGGATTAC
AGCATTGAAGAGGATGATTCGCAAACATTTGACATTATT

FIGURE 97

GCAACCTCAGCTTCTAGTATCCAGACTCCAGCGCCGCCCCGGGCGCGGACCCCAACCCCGAC
CCAGAGCTTCTCCAGCGGCGGCGCAGCGAGCAGGGCTCCCCGCCTTAACTTCCTCCGCGGGG
CCCAGCCACCTTCGGGAGTCCGGGTTGCCCACCTGCAAACTCTCCGCCTTCTGCACCTGCCA
CCCCTGAGCCAGCGCGGGCCCCCGAGCGAGTC**ATG**GCCAACGCGGGGCTGCAGCTGTTGGGC
TTCATTCTCGCCTTCCTGGGATGGATCGGCGCCATCGTCAGCACTGCCCTGCCCCAGTGGAG
GATTTACTCCTATGCCGGCGACAACATCGTGACCGCCCAGGCCATGTACGAGGGGCTGTGGA
TGTCCTGCGTGTCGCAGAGCACCGGGCAGATCCAGTGCAAAGTCTTTGACTCCTTGCTGAAT
CTGAGCAGCACATTGCAAGCAACCCGTGCCTTGATGGTGGTTGGCATCCTCCTGGGAGTGAT
AGCAATCTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGTGCTTGGAAGACGATGAGGTGC
AGAAGATGAGGATGGCTGTCATTGGGGGTGCGATATTTCTTCTTGCAGGTCTGGCTATTTTA
GTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATTCTATGACCCTATGACCCCAGT
CAATGCCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCC
TTCTGGGAGGTGCCCTACTTTGCTGTTCCTGTCCCCGAAAAACAACCTCTTACCCAACACCA
AGGCCCTATCCAAAACCTGCACCTTCCAGCGGGAAAGACTACGTG**TGA**CACAGAGGCAAAAG
GAGAAAATCATGTTGAAACAAACCGAAAATGGACATTGAGATACTATCATTAACATTAGGAC
CTTAGAATTTTGGGTATTGTAATCTGAAGTATGGTATTACAAAACAAACAAACAAACAAAAA
ACCCATGTGTTAAAATACTCAGTGCTAAACATGGCTTAATCTTATTTTATCTTCTTTCCTCA
ATATAGGAGGGAAGATTTTTCCATTTGTATTACTGCTTCCCATTGAGTAATCATACTCAAAT
GGGGGAAGGGGTGCTCCTTAAATATATATAGATATGTATATATACATGTTTTTCTATTAAAA
ATAGACAGTAAAATACTATTCTCATTATGTTGATACTAGCATACTTAAAATATCTCTAAAAT
AGGTAAATGTATTTAATTCCATATTGATGAAGATGTTTATTGGTATATTTTCTTTTTCGTCC
TTATATACATATGTAACAGTCAAATATCATTTACTCTTCTTCATTAGCTTTGGGTGCCTTTG
CCACAAGACCTAGCCTAATTTACCAAGGATGAATTCTTTCAATTCTTCATGCGTGCCCTTTT
CATATACTTATTTTATTTTTTACCATAATCTTATAGCACTTGCATCGTTATTAAGCCCTTAT
TTGTTTTGTGTTTCATTGGTCTCTATCTCCTGAATCTAACACATTTCATAGCCTACATTTTA
GTTTCTAAAGCCAAGAAGAATTTATTACAAATCAGAACTTTGGAGGCAAATCTTTCTGCATG
ACCAAAGTGATAAATTCCTGTTGACCTTCCCACACAATCCCTGTACTCTGACCCATAGCACT
CTTGTTTGCTTTGAAAATATTTGTCCAATTGAGTAGCTGCATGCTGTTCCCCCAGGTGTTGT
AACACAACTTTATTGATTGAATTTTTAAGCTACTTATTCATAGTTTTATATCCCCCTAAACT
ACCTTTTTGTTCCCCATTCCTTAATTGTATTGTTTTCCCAAGTGTAATTATCATGCGTTTTA
TATCTTCCTAATAAGGTGTGGTCTGTTTGTCTGAACAAAGTGCTAGACTTTCTGGAGTGATA
ATCTGGTGACAAATATTCTCTCTGTAGCTGTAAGCAAGTCACTTAATCTTTCTACCTCTTTT
TTCTATCTGCCAAATTGAGATAATGATACTTAACCAGTTAGAAGAGGTAGTGTGAATATTAA
TTAGTTTATATTACTCTTATTCTTTGAACATGAACTATGCCTATGTAGTGTCTTTATTTGCT
CAGCTGGCTGAGACACTGAAGAAGTCACTGAACAAAACCTACACACGTACCTTCATGTGATT
CACTGCCTTCCTCTCTCTACCAGTCTATTTCCACTGAACAAAACCTACACACATACCTTCAT
GTGGTTCAGTGCCTTCCTCTCTCTACCAGTCTATTTCCACTGAACAAAACCTACGCACATAC
CTTCATGTGGCTCAGTGCCTTCCTCTCTCTACCAGTCTATTTCCATTCTTTCAGCTGTGTCT
GACATGTTTGTGCTCTGTTCCATTTTAACAACTGCTCTTACTTTTCCAGTCTGTACAGAATG
CTATTTCACTTGAGCAAGATGATGTAATGGAAAGGGTGTTGGCACTGGTGTCTGGAGACCTG
GATTTGAGTCTTGGTGCTATCAATCACCGTCTGTGTTTGAGCAAGGCATTTGGCTGCTGTAA
GCTTATTGCTTCATCTGTAAGCGGTGGTTTGTAATTCCTGATCTTCCCACCTCACAGTGATG
TTGTGGGGATCCAGTGAGATAGAATACATGTAAGTGTGGTTTTGTAATTTAAAAAGTGCTAT
ACTAAGGGAAAGAATTGAGGAATTAACTGCATACGTTTTGGTGTTGCTTTTCAAATGTTTGA
AAATAAAAAAAATGTTAAG

FIGURE 98

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52185
><subunit 1 of 1, 211 aa, 1 stop
><MW: 22744, pI: 8.51, NX(S/T): 1
MANAGLQLLGFILAFLGWIGAIVSTALPQWRIYSYAGDNIVTAQAMYEGLWMSCVSQSTGQI
QCKVFDSLLNLSSTLQATRALMVVGILLGVIAIFVATVGMKCMKCLEDDEVQKMRMAVIGGA
IFLLAGLAILVATAWYGNRIVQEFYDPMTPVNARYEFGQALFTGWAAASLCLLGGALLCCSC
PRKTTSYPTPRPYPKPAPSSGKDYV

Important features:

Signal peptide:
amino acids 1-21

Transmembrane domains:
amino acids 82-102, 118-142 and 161-187

N-glycosylation site.
amino acids 72-75

PMP-22 / EMP / MP20 family proteins
amino acids 70-111

ABC-2 type transport system integral membrane protein
amino acids 119-133

FIGURE 99

TTCTGGCCAAACCCGGGGCTNCAGCTGTTGGGCTTCATCTCGCCTTCCTGGGATGGATCGGC
GCCATCNTCACACTGCCCTTCCCCAGTGGAGGATTTTACTCCCTATGCTGGCGACAACATCG
TGACCGCCCAGCCCATGTACGAGGGGCTGTGGATGTCCNGCGTGTCGCAGAGCACCGGGCAG
ATCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAGCAGCACATTGCAAGCAACCCGTGC
CTTGATGGTGGTTGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTTGGCATGA
AGTGTATGAAGTGCTTGGAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGGC
GCGATATTTCTTCTTGCAGGTCTGGCTATTTTAGTTGCCACAGCATGGTATGGCAATAGAAN
CNTTCAACANTTCTATGACCCTATGACCCCAGTCAATGCCAGGTACGAATTTGGTCA
GGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTGCCCTACTTTGCT
GTTCCTGTCCC

FIGURE 100

ACCCTTGACCCAACGCGGCCCCCCGACCGNTTCATGGCCAAACGCGGGNCTCCAGCTGTTGG
GCTTCATTCTCCCCTTCCTGGGATGGACCGGCGCCCATCNTCAGCACTGCCCTGCCCCAGTG
GAGGATTTACTCCTATNCCGGCNACAACATCGTGACCGCCCAGGCCNTGTACGAGGGGCTGT
GGATGTCCTGCGTGTCGCAGAGCACCGGGCAGATCCAGTGCAAAGTCTTTGACTCCCTTGCT
GAATCTGAGCAGCACATTGCAAGCAACCCGTGCCTTGATGGTGGTTGGCATCCTCCTGGGAG
TGATAGCAATCTTNNTGGCCACCGTTGTNNNTGAAGTGTATGAAGTGCTTGGAAGACGATGA
GGTGCAGAAGATGAGGATGGCTGTCATTGGGGGCGCGATATTTCTTCTTGCAGGTCTGGCTA
TTTTAGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATTCTATGACCCTATGACCGA

FIGURE 101

GGGCCCGACCATTATCCAACCGGGNTCACTGTTGGCTCATCTCCCTCCTGGATGAANCGCGC
CATCNTCAGACTCCCTGCCCCATGGAGATTTNNCCTATGCTGGCGACAACATCNTGACCCCC
AGCCATGTACGAGGGGCTTTGAACGTCNGCGTGTCGCAGANCACCGGGCAGATCCAGTGCAA
AGTCTTTGACTCCTTGCTGAATCTGNGCAGCACATTGCAGCAACCCNTGCCCTGATGGTGGT
TGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGT
GCTTGGAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGGCGCGATATTTCTT
CTTGCAGGTCTGGCTATTTNNNGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAAT
TCTATGACCCTATGACCCCAGTCAATGCCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGC
TGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTGCCCTACTTTGCTGTTCCTGCGA

FIGURE 102

ATTCTCCCCTCCTGGATGGATCGCNCCACCGTCACATTGCCTTCCCCCANTGGAGGATTNAC
TCCTATGCTGGCGACAACATCGTGACCCCCCAGGCCATTTACCGAGGGGCTTTGGATGTCNT
GCNTGTCGCAGAGCACCGGGCAGATCCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAG
CAGCACATTGCAAGCAACCCGTGCCTTGATGGGGTTGGCATCCTCCTGGGAGTGATAGCAAC
CTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGTGCTTGGAAGACGATGAGGTGCCAGAAG
ATGAGGATGGCTGTCATTGGGGGCGCGATATTTCTTGTTGCAGGTCTGGCTATTTTAGTNGC
CACAGCATGGTATGGCAATAGANTNNTTCNNGNNNTCTATGACCCTATGACCCCAGTCAATG
CCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTG
GGAGGTGCCCTACTTTGCTGTTCCTGTCCC

FIGURE 103

AGAGCACCGGCAGATCCCAGTNCAAAGTCTTTGACCCTTGCTGAATCTGAGCAGCACATTNC
AAGCAACCCCTTGCCTTGAAGGTGGTTGNCATCCCCCCTGGGAGTGAATAGCAATCTTTGTG
GCCACCGTTGGCATGAAGTNTATGAAGTGCTTGGAAGACGATGAGGTGCAGAAGATGAGGAT
GGCTGTCATTGGGGGCGCGATATTTCTTCTTGCAGGTCTGGCTATTTTAGTNNCCACAGCAT
GGTATGGCAATAGNATNNTTCGNGGNTTCTATGACCCTATGACCCCAGTCAATGCCAGGTAC
GAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTGC
CCTACTTTGCTGTTCCTGTCCCCGAA

FIGURE 104

AGCAATGCCCTGCCCCCAGTGGAGGATTAATTCCTATGNTGGGGACAACATTGTGACNGCCC
AGGCCATGTACGGGGGGCTGTGGATGTCCTGCGTGTCGCAGAGCACCGGGCAGATCCAGTGC
AAAGTNTTTGACTCCTTGCTGAATTTGAGCAGCACATTGCAAGCAACCCGTGCCTTGATGGT
GGTTGGCATCTTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTGGNAATGAAGTGTATGA
AGTGCTTGGAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGGCGCGATATTT
CTTNTTGCAGGTCTGGCTATTTTAGTTGCCACAGCATGGTATGGCAATAGAATNGTTCAAGA
ATTTTATGACCCTATGACCCCAGTCAATGCCAGGTACGAATTTGGTCAGGCTTTNTTCACTG
GCTGGGCTGCTGCTTNTTTCTGCCTTNTGGGAGGTGCCCTANTTTGCTGTTCCTGCGAACC

FIGURE 105

TCATAGGGGGGCGCGATATTTTTTCTTGCAGGTNTGGTTATTTTAGTTGCCACAGCATGGTA
TGGCAATAGAATCGTTCAAGAATTNTATGACCCTATGACCCCAGTCAATGCCAGGTACGAAT
TTGGTCAGGCTCTNTTCACTGGNTGGGCTGCTGCTTCTNTNNGCCTTNTGGGAGGTGCCCTA
CTTTGCTGTTCCTG

FIGURE 106

TTCCTGGGATGGATCCGCCCCCATCNTCACATGCCCTGCCCCNTGGAGATTTACNCCTATGC
TGGCGAACAACATCNTGACCGCCCAGGCCATGTACGAGGGGCTGTGGAATGTCCTGCGTGTC
CCAGAGCACCGGGCAGATCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAGCAGCACAT
TGCAAGCAACCNTGCCTTGATGGTGGTTGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGG
CCACCGTTGGCATGAAAGTGTATGAAGTGCTTGGAAGACGATGAGGTGCAGAAGATGAGGAT
GGCTGTCATTGGGGGCGCGATATTTCTTCTTGCAGGTCTGGCTATTTTAGNNGCCACAGCAT
GGTATGGCAATCAGACCCNNTCANAAACTCTATGACCCTATGACCCCAGTCAATGCCAGGTA
CGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTG
CCCTACTTTGCTGTTCCTGTCCCCGAAAAACAACCTCTTACCCACG

FIGURE 107

CGGGGCTGCAGCTGTTGGGCTTCATCTCGCTTCCTGGGATGGAATCGGCGCCATCGTCAGCA
CTGCCCTGCCCCATGGAGGATTTACTCNTATGCTGGCGACAACATCGTGACCNCCCAGGCCA
TGTACGAGGGGCTGTGGATGTCNGCGTGTCGCAGAGCACCGGGCAGATCCAGTGCAAAGTCT
TTGACTCCTTGCTGAATCTGAGCAGCACATTGCAAGCAACCNTGCCTTGATGGTGGTTGGCA
TCCTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGTGCTTG
GAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGGCGCGATATTTCTTCTTGC
AGGTCTGGCTATTTNTAGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATTCTAT
GACCCTATGACCCCAGTCAATGCCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGCTGGGC
TGCTGCTTCTCTCTGCCTTCTGGGAGGTGCCCTACTTTGCTGTTCCTGCGAA

FIGURE 108

GCGTGCCGTCAGCTCGCCGGGCACCGCGGCCTCGCCCTCGCCCTCCGCCCCTGCGCCTGCAC
CGCGTAGACCGACCCCCCCCTCCAGCGCGCCCACCCGGTAGAGGACCCCCGCCCGTGCCCCG
ACCGGTCCCCGCCTTTTTGTAAAACTTAAAGCGGGCGCAGCATTAACGCTTCCCGCCCCGGT
GACCTCTCAGGGGTCTCCCCGCCAAAGGTGCTCCGCCGCTAAGGAAC**ATG**GCGAAGGTGGAG
CAGGTCCTGAGCCTCGAGCCGCAGCACGAGCTCAAATTCCGAGGTCCCTTCACCGATGTTGT
CACCACCAACCTAAAGCTTGGCAACCCGACAGACCGAAATGTGTGTTTTAAGGTGAAGACTA
CAGCACCACGTAGGTACTGTGTGAGGCCCAACAGCGGAATCATCGATGCAGGGGCCTCAATT
AATGTATCTGTGATGTTACAGCCTTTCGATTATGATCCCAATGAGAAAAGTAAACACAAGTT
TATGGTTCAGTCTATGTTTGCTCCAACTGACACTTCAGATATGGAAGCAGTATGGAAGGAGG
CAAAACCGGAAGACCTTATGGATTCAAAACTTAGATGTGTGTTTGAATTGCCAGCAGAGAAT
GATAAACCACATGATGTAGAAATAAATAAAATTATATCCACAACTGCATCAAAGACAGAAAC
ACCAATAGTGTCTAAGTCTCTGAGTTCTTCTTTGGATGACACCGAAGTTAAGAAGGTTATGG
AAGAATGTAAGAGGCTGCAAGGTGAAGTTCAGAGGCTACGGGAGGAGAACAAGCAGTTCAAG
GAAGAAGATGGACTGCGGATGAGGAAGACAGTGCAGAGCAACAGCCCCATTTCAGCATTAGC
CCCAACTGGGAAGGAAGAAGGCCTTAGCACCCGGCTCTTGGCTCTGGTGGTTTTGTTCTTTA
TCGTTGGTGTAATTATTGGGAAGATTGCCTTG**TAG**AGGTAGCATGCACAGGATGGTAAATTG
GATTGGTGGATCCACCATATCATGGGATTTAAATTTATCATAACCATGTGTAAAAAGAAATT
AATGTATGATGACATCTCACAGGTCTTGCCTTTAAATTACCCCTCCCTGCACACACATACAC
AGATACACACACACAAATATAATGTAACGATCTTTTAGAAAGTTAAAAATGTATAGTAACTG
ATTGAGGGGGAAAAAGAATGATCTTTATTAATGACAAGGGAAACCATGAGTAATGCCACAAT
GGCATATTGTAAATGTCATTTTAAACATTGGTAGGCCTTGGTACATGATGCTGGATTACCTC
TCTTAAAATGACACCCTTCCTCGCCTGTTGGTGCTGGCCCTTGGGGAGCTGGAGCCCAGCAT
GCTGGGGAGTGCGGTCAGCTCCACACAGTAGTCCCCACGTGGCCCACTCCCGGCCCAGGCTG
CTTTCCGTGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGACTGATGAACAGAGTCAGA
AGCCCAAAGGAATTGCACTGTGGCAGCATCAGACGTACTCGTCATAAGTGAGAGGCGTGTGT
TGACTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTCACTTAAAGGGACCAA
GCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAACTGTTATTCAGAGATGTTTAATGCATA
TTTAACTTATTTAATGTATTTCATCTCATGTTTTCTTATTGTCACAAGAGTACAGTTAATGC
TGCGTGCTGCTGAACTCTGTTGGGTGAACTGGTATTGCTGCTGGAGGGCTGTGGGCTCCTCT
GTCTCTGGAGAGTCTGGTCATGTGGAGGTGGGGTTTATTGGGATGCTGGAGAAGAGCTGCCA
GGAAGTGTTTTTTCTGGGTCAGTAAATAACAACTGTCATAGGGAGGGAAATTCTCAGTAGTG
ACAGTCAACTCTAGGTTACCTTTTTTAATGAAGAGTAGTCAGTCTTCTAGATTGTTCTTATA
CCACCTCTCAACCATTACTCACACTTCCAGCGCCCAGGTCCAAGTCTGAGCCTGACCTCCCC
TTGGGGACCTAGCCTGGAGTCAGGACAAATGGATCGGGCTGCAGAGGGTTAGAAGCGAGGGC
ACCAGCAGTTGTGGGTGGGGAGCAAGGGAAGAGAGAAACTCTTCAGCGAATCCTTCTAGTAC
TAGTTGAGAGTTTGACTGTGAATTAATTTTATGCCATAAAAGACCAACCCAGTTCTGTTTGA
CTATGTAGCATCTTGAAAAGAAAAATTATAATAAAGCCCCAAAATTAAGAAAA

FIGURE 109

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53977

<subunit 1 of 1, 243 aa, 1 stop

<MW: 27228, pI: 7.43, NX(S/T): 2

MAKVEQVLSLEPQHELKFRGPFTDVVTTNLKLGNPTDRNVCFKVKTTAPRRYCVRPNSGIID
AGASINVSVMLQPFDYDPNEKSKHKFMVQSMFAPTDTSDMEAVWKEAKPEDLMDSKLRCVFE
LPAENDKPHDVEINKIISTTASKTETPIVSKSLSSSLDDTEVKKVMEECKRLQGEVQRLREE
NKQFKEEDGLRMRKTVQSNSPISALAPTGKEEGLSTRLLALVVLFFIVGVIIGKIAL

Important features:

Putative transmembrane domain:

amino acids 224-239

N-glycosylation site.

amino acids 68-71

N-myristoylation site.

amino acids 59-64, 64-69 and 235-240

FIGURE 110

GTCAGTCTTCTAGATTGTCCTTATCCCACCTTTCAACCANTACTCACATTTCNAGCGCCCAG
GTCCANGTCTGAGCCTGACTTCCCCTTGGGGACCTAGCCTGGAGTCAGGACAATGGNTCGGG
CTGCAGAGGNTTAGAAGCGAGGGCACCAGCAGTTTTGGGTGGGGAGCAAGGGNNGAGAGAAA
CTCTTCAGCGAATCCTTCTAGTACTAGTTGAGAGTTTGACTGTGAATTAATTTTATGCCATA
AAAGACNAACCCAGTTCTGTTTGACTATGTAGCATCTTGAAAAGAAAAATTATAATAAAGCC
CCAAAATTAAGAATTCTTTTGTCATTTTGTCACATTTGCTCTATGGGGGGAATTATTATTTT
ATCATTTTTATTATTTTGCCATTGGAAGGTTAACTTTAAAATGAGC

FIGURE 111

TATTGTAAAGGCCATTTTAAACCATTGGTAGGCCTTGGTACATGATGCTGGATTACCTCCTT
AAATGACACCNTTCCTCGCCTGTTGGTGCTGGCCNTTGGGGAGCTGGAGCCCCAGCATGCTG
GGGAGTGCGGTCAGCTCCACACAGTAGTCCCCACGTGGCCCACTCCCGGCCCAGGCTGCTTT
CCGTGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGACTGATGAACAGAGTCAGAAGCC
CAAAGGAATTGCCACTGTGGCAGCATCAGACGTACTCGTCATAAGTGAGAGGCGTGTGTTGA
CTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTCACTTAAAGGGACCAAGCT
AAATTGTATTGGTTCATGTAGTGAAGTCAAACTGTTATTCAGAGATGTTTAATGCATATTTA
ACTTATTTAATGTATTTCATCTCATGTTTTCTTATTGTCACAAGAGTACAGTTAATGCTGCG
TGCTGCTGAACTCTGTTGGGTGAACTGGTATTGCTGCTGGAGGGCTG

FIGURE 112

CCCTGGTGGTTTTGTTCTTTAATTCGTTGGTGTAATTNTTGGGAAGATTGCTTGTAGAGGTA
GNATGCACCNGGCTGGTAAATTGGATTGGTGGATCCACCATATCCATGGGATTTAAATTTAT
CATAACCATGTGTAAAAAGAAATTAATGTATGATGACATNTCACAGGTATTGCCTTTAAATT
ACCCATCCCTGNANACACATACACAGATACACANANACAAATNTAATGTAACGATNTTTTAG
AAAGTTAAAAATGTATAGTAAC

FIGURE 113

GGTGGCCCATTCCCGGCCCAGGCTGCTTTCCGGTNTTCAGTTCTGTCCAAGCCATCAGCTCC
TTGGGACTGATGAACAGAGTCAGAAGCCCAAAGGAATTGCACTGTGGCAGCATNAGACGTAC
TTGTNATAAGTGAGAGGCGTGTGTTGACTGATTGACCCAGCGCTTTGGAAATAAATGGCAGT
GCTTTGTTCANTTAAAGGGACCAAGCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAACTG
TTATTCAGAGATGTTTAATGCATATTTAANTTATTTAATGTATTTNATNTCATGTTTTCTTA
TTGTCACAAGAGTACAGTTAATGCTGCGTGCTGCTGAANTNTGTTGGGTGAACTGGTATTGC
TGCTGGAGGGCTGTGGGCTCCTCTGTCTTTGGAGAGTCTGGTCATGTGGAGGTGGG

FIGURE 114

TGCTTTCCGTGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGACTTGATGAACAGAGTC
AGAAGCCCAAAGGAATTGCACTGTGGCAGCATCAGACGTACTCGTCATAAGTGAGAGGCGTG
TGTTGACTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTCACTTAAAGGGAC
CAAGCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAACTGTTATTCAGAGATGTTTAATGC
ATATTTAACTTATTTAATGTATTTCATCTCATGTTTTCTTATTGTCACAAGAGTACAGTTAA
TGCTGCGTGC

FIGURE 115

AAACCTTTAAAAGTTGAGGGGAAAAGAATGATCCTTTATTAATGACAAGGGAAACCNTGNGT
AATGCCACAATGGCATATTGTAAATGTCATTTTAAACATTGGTAGGCCTTGGTACATGATGC
TGGATTACCTCTCTTAAAATGACACCCTTCCTCGCCTGTTGGTGCTGGCCCTTGGGGAGCTN
GAGCCCAGCATGCTGGGGAGTGCGGTCTGCTCCACACAGTAGTCCCCANGTGGCCCANTCCC
GGCCCAGGCTGCTTTCCGTGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGANTGATGA
ACAGAGTCAGAAGCCCAAAGGAATTGCANTGTGGCAGCATCAGANGTANTNGTCATAAGTGA
GAGGCGTGTGTTGANTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTCANTT
AAAGGGNCCAAGNTAAATTTGTATTGGTTCATGTAGTGAAGTCAAANTGTTATTCAGAGATG
TTTAATGCATATTTAANTTATTTAATGTATTTCATNTCATGTTTTCTTATTGTCACAAGGGT
ACAGTTAATGCTGCGTGCTGCTGAANTCTGTTGGGTGAANTGGTATTGCTG

FIGURE 116

GGCCCTTGGGGAGCTGGAGCCCAGCATGCTGGGGAGTGCGGTCAGCTCCACACAGTAGTCCC
CACGTGGCCCACTCCCGGCCCAGGCTGCTTTCCGTGTCTTCAGTTCTGTCCAAGCCATCAGC
TCCTTGGGACTGATGAACAGAGTCAGAAGCCCAAAGGAATTGCACTGTGGCAGCATCAGACG
TACTCGTCATAAGTGAGAGGCGTGTGTTGACTGATTGACCCAGCGCTTTGGAAATAAATGGC
AGTGCTTTGTTCACTTAAAGGGACCAAGCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAA
CTGTTATTCAGAGATGTTTAATGCATATTTAACTTATTTAATGTATTTCATCTCATGTTTTC
TTATTGTCACAAGAGTACAGTTAATGCTGCGTGCTGCTGAACTCTGTTGGGTGAACTGGTAT
TGCTGCTGGAGGGCTGTGGGCTCCTCTGTCTCTGGAGAGTCTGGTCATGTGGAGGTGGG

FIGURE 117

GCGAGCTCCGGGTGCTGTGGCCCGGCCTTGGCGGGGCGGCCTCCGGCTCAGGCTGGCTGAGA
GGCTCCCAGCTGCAGCGTCCCCGCCCGCCTCCTCGGGAGCTCTGATCTCAGCTGACAGTGCC
CTCGGGGACCAAACAAGCCTGGCAGGGTCTCACTTTGTTGCCCAGGCTGGAGTTCAGTGCCA
TGATCATGGTTTACTGCAGCCTTGACCTCCTGGGTTCAAGCGATCCTGCTGAGTAGCTGGGA
CTACAGGACAAAATTAGAAGATCAAA**ATG**GAAAATATGCTGCTTTGGTTGATATTTTCACC
CCTGGGTGGACCCTCATTGATGGATCTGAAATGGAATGGGATTTTATGTGGCACTTGAGAAA
GGTACCCCGGATTGTCAGTGAAAGGACTTTCCATCTCACCAGCCCCGCATTTGAGGCAGATG
CTAAGATGATGGTAAATACAGTGTGTGGCATCGAATGCCAGAAAGAACTCCCAACTCCCAGC
CTTTCTGAATTGGAGGATTATCTTTCCTATGAGACTGTCTTTGAGAATGGCACCCGAACCTT
AACCAGGGTGAAAGTTCAAGATTTGGTTCTTGAGCCGACTCAAAATATCACCACAAAGGGAG
TATCTGTTAGGAGAAAGAGACAGGTGTATGGCACCGACAGCAGGTTCAGCATCTTGGACAAA
AGGTTCTTAACCAATTTCCCTTTCAGCACAGCTGTGAAGCTTTCCACGGGCTGTAGTGGCAT
TCTCATTTCCCCTCAGCATGTTCTAACTGCTGCCCACTGTGTTCATGATGGAAAGGACTATG
TCAAAGGGAGTAAAAAGCTAAGGGTAGGGTTGTTGAAGATGAGGAATAAAAGTGGAGGCAAG
AAACGTCGAGGTTCTAAGAGGAGCAGGAGAGAAGCTAGTGGTGGTGACCAAAGAGAGGGTAC
CAGAGAGCATCTGCAGGAGAGAGCGAAGGGTGGGAGAAGAAGAAAAAAATCTGGCCGGGGTC
AGAGGATTGCCGAAGGGAGGCCTTCCTTTCAGTGGACCCGGGTCAAGAATACCCACATTCCG
AAGGGCTGGGCACGAGGAGGCATGGGGGACGCTACCTTGGACTATGACTATGCTCTTCTGGA
GCTGAAGCGTGCTCACAAAAAGAAATACATGGAACTTGGAATCAGCCCAACGATCAAGAAAA
TGCCTGGTGGAATGATCCACTTCTCAGGATTTGATAACGATAGGGCTGATCAGTTGGTCTAT
CGGTTTTGCAGTGTGTCCGACGAATCCAATGATCTCCTTTACCAATACTGCGATGCTGAGTC
GGGCTCCACCGGTTCGGGGGTCTATCTGCGTCTGAAAGATCCAGACAAAAAGAATTGGAAGC
GCAAAATCATTGCGGTCTACTCAGGGCACCAGTGGGTGGATGTCCACGGGGTTCAGAAGGAC
TACAACGTTGCTGTTCGCATCACTCCCCTAAAATACGCCCAGATTTGCCTCTGGATTCACGG
GAACGATGCCAATTGTGCTTACGGC**TAA**CAGAGACCTGAAACAGGGCGGTGTATCATCTAAA
TCACAGAGAAAACCAGCTCTGCTTACCGTAGTGAGATCACTTCATAGGTTATGCCTGGACTT
GAACTCTGTCAATAGCATTTCAACATTTTTCAAAATCAGGAGATTTTCGTCCATTTAAAAAA
TGTATAGGTGCAGATATTGAAACTAGGTGGGCACTTCAATGCCAAGTATATACTCTTCTTTA
CATGGTGATGAGTTTCATTTGTAGAAAAATTTTGTTGCCTTCTTAAAAATTAGACACACTTT
AAACCTTCAAACAGGTATTATAAATAACATGTGACTCCTTAATGGACTTATTCTCAGGGTCC
TACTCTAAGAAGAATCTAATAGGATGCTGGTTGTGTATTAAATGTGAAATTGCATAGATAAA
GGTAGATGGTAAAGCAATTAGTATCAGAATAGAGACAGAAAGTTACAACACAGTTTGTACTA
CTCTGAGATGGATCCATTCAGCTCATGCCCTCAATGTTTATATTGTGTTATCTGTTGGGTCT
GGGACATTTAGTTTAGTTTTTTTGAAGAATTACAAATCAGAAGAAAAAGCAAGCATTATAAA
CAAAACTAATAACTGTTTTACTGCTTTAAGAAATAACAATTACAATGTGTATTATTTAAAAA
TGGGAGAAATAGTTTGTTCTATGAAATAAACCTAGTTTAGAAATAGGGAAGCTGAGACATTT
TAAGATCTCAAGTTTTTATTTAACTAATACTCAAAATATGGACTTTTCATGTATGCATAGGG
AAGACACTTCACAAATTATGAATGATCATGTGTTGAAAGCCACATTATTTTATGCTATACAT
TCTATGTATGAGGTGCTACATTTTTAGGACAAAGAATTCTGTAATCTTTTTCAAGAAAGAGT
CTTTTTCTCCTTGACAAAATCCAGCTTTTGTATGAGGACTATAGGGTGAATTCTCTGATTAG
TAATTTTAGATATGTCCTTTCCTAAAAATGAATAAAATTTATGAATATGA

FIGURE 118

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57253

<subunit 1 of 1, 413 aa, 1 stop

<MW: 47070, pI: 9.92, NX(S/T): 3

MENMLLWLIFFTPGWTLIDGSEMEWDFMWHLRKVPRIVSERTFHLTSPAFEADAKMMVNTVC
GIECQKELPTPSLSELEDYLSYETVFENGTRTLTRVKVQDLVLEPTQNITTKGVSVRRKRQV
YGTDSRFSILDKRFLTNFPFSTAVKLSTGCSGILISPQHVLTAAHCVHDGKDYVKGSKKLRV
GLLKMRNKSGGKKRRGSKRSRREASGGDQREGTREHLQERAKGGRRRKKSGRGQRIAEGRPS
FQWTRVKNTHIPKGWARGGMGDATLDYDYALLELKRAHKKKYMELGISPTIKKMPGGMIHFS
GFDNDRADQLVYRFCSVSDESNDLLYQYCDAESGSTGSGVYLRLKDPDKKNWKRKIIAVYSG
HQWVDVHGVQKDYNVAVRITPLKYAQICLWIHGNDANCAYG

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 90-93, 110-113 and 193-196

Glycosaminoglycan attachment site.

amino acids 236-239

Serine proteases, trypsin family, histidine active site.

amino acids 165-170

FIGURE 119

AATGTGAGAGGGGCTGATGGAAGCTGATAGGCAGGACTGGAGTGTTAGCACCAGTACTGGAT
GTGACAGCAGGCAGAGGAGCACTTAGCAGCTTATTCAGTGTCCGATTCTGATTCCGGCAAGG
ATCCAAGC**ATG**GAATGCTGCCGTCGGGCAACTCCTGGCACACTGCTCCTCTTTCTGGCTTTC
CTGCTCCTGAGTTCCAGGACCGCACGCTCCGAGGAGGACCGGGACGGCCTATGGGATGCCTG
GGGCCCATGGAGTGAATGCTCACGCACCTGCGGGGGAGGGGCCTCCTACTCTCTGAGGCGCT
GCCTGAGCAGCAAGAGCTGTGAAGGAAGAAATATCCGATACAGAACATGCAGTAATGTGGAC
TGCCCACCAGAAGCAGGTGATTTCCGAGCTCAGCAATGCTCAGCTCATAATGATGTCAAGCA
CCATGGCCAGTTTTATGAATGGCTTCCTGTGTCTAATGACCCTGACAACCCATGTTCACTCA
AGTGCCAAGCCAAAGGAACAACCCTGGTTGTTGAACTAGCACCTAAGGTCTTAGATGGTACG
CGTTGCTATACAGAATCTTTGGATATGTGCATCAGTGGTTTATGCCAAATTGTTGGCTGCGA
TCACCAGCTGGGAAGCACCGTCAAGGAAGATAACTGTGGGGTCTGCAACGGAGATGGGTCCA
CCTGCCGGCTGGTCCGAGGGCAGTATAAATCCCAGCTCTCCGCAACCAAATCGGATGATACT
GTGGTTGCACTTCCCTATGGAAGTAGACATATTCGCCTTGTCTTAAAAGGTCCTGATCACTT
ATATCTGGAAACCAAAACCCTCCAGGGGACTAAAGGTGAAAACAGTCTCAGCTCCACAGGAA
CTTTCCTTGTGGACAATTCTAGTGTGGACTTCCAGAAATTTCCAGACAAAGAGATACTGAGA
ATGGCTGGACCACTCACAGCAGATTTCATTGTCAAGATTCGTAACTCGGGCTCCGCTGACAG
TACAGTCCAGTTCATCTTCTATCAACCCATCATCCACCGATGGAGGGAGACGGATTTCTTTC
CTTGCTCAGCAACCTGTGGAGGAGGTTATCAGCTGACATCGGCTGAGTGCTACGATCTGAGG
AGCAACCGTGTGGTTGCTGACCAATACTGTCACTATTACCCAGAGAACATCAAACCCAAACC
CAAGCTTCAGGAGTGCAACTTGGATCCTTGTCCAGCCAGTGACGGATACAAGCAGATCATGC
CTTATGACCTCTACCATCCCCTTCCTCGGTGGGAGGCCACCCCATGGACCGCGTGCTCCTCC
TCGTGTGGGGGGGGCATCCAGAGCCGGGCAGTTTCCTGTGTGGAGGAGGACATCCAGGGGCA
TGTCACTTCAGTGGAAGAGTGGAAATGCATGTACACCCCTAAGATGCCCATCGCGCAGCCCT
GCAACATTTTTGACTGCCCTAAATGGCTGGCACAGGAGTGGTCTCCGTGCACAGTGACATGT
GGCCAGGGCCTCAGATACCGTGTGGTCCTCTGCATCGACCATCGAGGAATGCACACAGGAGG
CTGTAGCCCAAAAACAAAGCCCCACATAAAAGAGGAATGCATCGTACCCACTCCCTGCTATA
AACCCAAAGAGAAACTTCCAGTCGAGGCCAAGTTGCCATGGTTCAAACAAGCTCAAGAGCTA
GAAGAAGGAGCTGCTGTGTCAGAGGAGCCCTCG**TAA**GTTGTAAAAGCACAGACTGTTCTATA
TTTGAAACTGTTTTGTTTAAAGAAAGCAGTGTCTCACTGGTTGTAGCTTTCATGGGTTCTGA
ACTAAGTGTAATCATCTCACCAAAGCTTTTTGGCTCTCAAATTAAAGATTGATTAGTTTCAA
AAAAAAAAA

FIGURE 120

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58847

<subunit 1 of 1, 525 aa, 1 stop

<MW: 58416, pI: 6.62, NX(S/T): 1

MECCRRATPGTLLLFLAFLLLSSRTARSEEDRDGLWDAWGPWSECSRTCGGGASYSLRRCLS
SKSCEGRNIRYRTCSNVDCPPEAGDFRAQQCSAHNDVKHHGQFYEWLPVSNDPDNPCSLKCQ
AKGTTLVVELAPKVLDGTRCYTESLDMCISGLCQIVGCDHQLGSTVKEDNCGVCNGDGSTCR
LVRGQYKSQLSATKSDDTVVALPYGSRHIRLVLKGPDHLYLETKTLQGTKGENSLSSTGTFL
VDNSSVDFQKFPDKEILRMAGPLTADFIVKIRNSGSADSTVQFIFYQPIIHRWRETDFFPCS
ATCGGGYQLTSAECYDLRSNRVVADQYCHYYPENIKPKPKLQECNLDPCPASDGYKQIMPYD
LYHPLPRWEATPWTACSSSCGGGIQSRAVSCVEEDIQGHVTSVEEWKCMYTPKMPIAQPCNI
FDCPKWLAQEWSPCTVTCGQGLRYRVVLCIDHRGMHTGGCSPKTKPHIKEECIVPTPCYKPK
EKLPVEAKLPWFKQAQELEEGAAVSEEPS

Important features:

Signal peptide:

amino acids 1-25

N-glycosylation site.

amino acids 251-254

Thrombospondin 1

amino acids 385-399

von Willebrand factor type C domain proteins

amino acids 385-399, 445-459 and 42-56

FIGURE 121

CGGACGCGTGGGCGGCGGCTGCGGAACTCCCGTGGAGGGGCCGGTGGGCCCTCGGGCCTGAC

AGATGGCAGTGGCCACTGCGGCGGCAGTACTGGCCGCTCTGGGCGGGGCGCTGTGGCTGGCG

GCCCGCCGGTTCGTGGGGCCCAGGGTCCAGCGGCTGCGCAGAGGCGGGGACCCCGGCCTCAT

GCACGGGAAGACTGTGCTGATCACCGGGGCGAACAGCGGCCTGGGCCGCGCCACGGCCGCCG

AGCTACTGCGCCTGGGAGCGCGGGTGATCATGGGCTGCCGGGACCGCGCGCGCGCCGAGGAG

GCGGCGGGTCAGCTCCGCCGCGAGCTCCGCCAGGCCGCGGAGTGCGGCCCAGAGCCTGGCGT

CAGCGGGGTGGGCGAGCTCATAGTCCGGGAGCTGGACCTCGCCTCGCTGCGCTCGGTGCGCG

CCTTCTGCCAGGAAATGCTCCAGGAAGAGCCTAGGCTGGATGTCTTGATCAATAACGCAGGG

ATCTTCCAGTGCCCTTACATGAAGACTGAAGATGGGTTTGAGATGCAGTTCGGAGTGAACCA

TCTGGGGCACTTTCTACTCACCAATCTTCTCCTTGGACTCCTCAAAAGTTCAGCTCCCAGCA

GGATTGTGGTAGTTTCTTCCAAACTTTATAAATACGGAGACATCAATTTTGATGACTTGAAC

AGTGAACAAAGCTATAATAAAAGCTTTTGTTATAGCCGGAGCAAACTGGCTAACATTCTTTT

TACCAGGGAACTAGCCCGCCGCTTAGAAGGCACAAATGTCACCGTCAATGTGTTGCATCCTG

GTATTGTACGGACAAATCTGGGGAGGCACATACACATTCCACTGTTGGTCAAACCACTCTTC

AATTTGGTGTCATGGGCTTTTTTCAAAACTCCAGTAGAAGGTGCCCAGACTTCCATTTATTT

GGCCTCTTCACCTGAGGTAGAAGGAGTGTCAGGAAGATACTTTGGGGATTGTAAAGAGGAAG

AACTGTTGCCCAAAGCTATGGATGAATCTGTTGCAAGAAAACTCTGGGATATCAGTGAAGTG

ATGGTTGGCCTGCTAAAATAGGAACAAGGAGTAAAAGAGCTGTTTATAAAACTGCATATCAG

TTATATCTGTGATCAGGAATGGTGTGGATTGAGAACTTGTTACTTGAAGAAAAAGAATTTTG

ATATTGGAATAGCCTGCTAAGAGGTACATGTGGGTATTTTGGAGTTACTGAAAAATTATTTT

TGGGATAAGAGAATTTCAGCAAAGATGTTTTAAATATATATAGTAAGTATAATGAATAATAA

GTACAATGAAAAATACAATTATATTGTAAAATTATAACTGGGCAAGCATGGATGACATATTA

ATATTTGTCAGAATTAAGTGACTCAAAGTGCTATCGAGAGGTTTTTCAAGTATCTTTGAGTT

TCATGGCCAAAGTGTTAACTAGTTTTACTACAATGTTTGGTGTTTGTGTGGAAATTATCTGC

CTGGTGTGTGCACACAAGTCTTACTTGGAATAAATTTACTGGTAC

FIGURE 122

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58747

<subunit 1 of 1, 336 aa, 1 stop

<MW: 36865, pI: 9.15, NX(S/T): 2

MAVATAAAVLAALGGALWLAARRFVGPRVQRLRRGGDPGLMHGKTVLITGANSGLGRATAAE
LLRLGARVIMGCRDRARAEEAAGQLRRELRQAAECGPEPGVSGVGELIVRELDLASLRSVRA
FCQEMLQEEPRLDVLINNAGIFQCPYMKTEDGFEMQFGVNHLGHFLLTNLLLGLLKSSAPSR
IVVVSSKLYKYGDINFDDLNSEQSYNKSFCYSRSKLANILFTRELARRLEGTNVTVNVLHPG
IVRTNLGRHIHIPLLVKPLFNLVSWAFFKTPVEGAQTSIYLASSPEVEGVSGRYFGDCKEEE
LLPKAMDESVARKLWDISEVMVGLLK

Important features:

Signal peptide:

amino acids 1-21

Short-chain alcohol dehydrogenase family protein

amino acids 134-144, 44-56 and 239-248

N-glycosylation site.

amino acids 212-215 and 239-242

FIGURE 123

GGGGATTGTAAAGAGGAAGNACTGTGCCCAAAGNTATGGATGAATCTGTTGCAAGAAAATTN
TGGGATATCAGTGAAGTGATGGTTNGCCTGCTAAAATAGGAACAAGGAGTAAAAGAGCTGTT
TATAAAACTGCATATCAGTTATATCTGTGATCAGGAATGGTGTGGATTGAGAACTTGTTACT
TGAAGAAAAAGAATTTTGATATTGGAATAGCCTGNTAAGAGGNACATGTGGGTATTTTGGAG
TTACTGAAAAATTATTTTTGGGATAAGAGAATTTCAGCAAAGATGTTTTAAATATATATAGT
AAGTATAATGAATAATAAGTACAATGAAAAATACAATTATATTGTAAAATTATAACTGGGCA
AGCATGGATGACATATTAATATTTGTCAGAATTAAGTGACTCAAAGTGCTATCGAGAGGTTT
TTCAAGTATCTTTGAGTTTCATGGCCAAAGTGTTAACTAGTTTTACTACAATGTTTGGTGTT
TGTGTGGAAATTATCTGCCTGGCTT

FIGURE 124

GAGAGGACGAGGTGCCGCTGCCTGGAGAATCCTCCGCTGCCGTCGGCTCCCGGAGCCCAGCC
CTTTCCTAACCCAACCCAACCTAGCCCAGTCCCAGCCGCCAGCGCCTGTCCCTGTCACGGAC
CCCAGCGTTACC**ATG**CATCCTGCCGTCTTCCTATCCTTACCCGACCTCAGATGCTCCCTTCT
GCTCCTGGTAACTTGGGTTTTTACTCCTGTAACAACTGAAATAACAAGTCTTGCTACAGAGA
ATATAGATGAAATTTTAAACAATGCTGATGTTGCTTTAGTAAATTTTTATGCTGACTGGTGT
CGTTTCAGTCAGATGTTGCATCCAATTTTTGAGGAAGCTTCCGATGTCATTAAGGAAGAATT
TCCAAATGAAAATCAAGTAGTGTTTGCCAGAGTTGATTGTGATCAGCACTCTGACATAGCCC
AGAGATACAGGATAAGCAAATACCCAACCCTCAAATTGTTTCGTAATGGGATGATGATGAAG
AGAGAATACAGGGGTCAGCGATCAGTGAAAGCATTGGCAGATTACATCAGGCAACAAAAAAG
TGACCCCATTCAAGAAATTCGGGACTTAGCAGAAATCACCACTCTTGATCGCAGCAAAAGAA
ATATCATTGGATATTTTGAGCAAAAGGACTCGGACAACTATAGAGTTTTTGAACGAGTAGCG
AATATTTTGCATGATGACTGTGCCTTTCTTTCTGCATTTGGGGATGTTTCAAAACCGGAAAG
ATATAGTGGCGACAACATAATCTACAAACCACCAGGGCATTCTGCTCCGGATATGGTGTACT
TGGGAGCTATGACAAATTTTGATGTGACTTACAATTGGATTCAAGATAAATGTGTTCCTCTT
GTCCGAGAAATAACATTTGAAAATGGAGAGGAATTGACAGAAGAAGGACTGCCTTTTCTCAT
ACTCTTTCACATGAAAGAAGATACAGAAAGTTTAGAAATATTCCAGAATGAAGTAGCTCGGC
AATTAATAAGTGAAAAAGGTACAATAAACTTTTTACATGCCGATTGTGACAAATTTAGACAT
CCTCTTCTGCACATACAGAAAACTCCAGCAGATTGTCCTGTAATCGCTATTGACAGCTTTAG
GCATATGTATGTGTTTGGAGACTTCAAAGATGTATTAATTCCTGGAAAACTCAAGCAATTCG
TATTTGACTTACATTCTGGAAAACTGCACAGAGAATTCCATCATGGACCTGACCCAACTGAT
ACAGCCCCAGGAGAGCAAGCCCAAGATGTAGCAAGCAGTCCACCTGAGAGCTCCTTCCAGAA
ACTAGCACCCAGTGAATATAGGTATACTCTATTGAGGGATCGAGATGAGCTT**TAA**AAACTTG
AAAAACAGTTTGTAAGCCTTTCAACAGCAGCATCAACCTACGTGGTGGAAATAGTAAACCTA
TATTTTCATAATTCTATGTGTATTTTTATTTTGAATAAACAGAAAGAAATTTAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 125

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57689

<subunit 1 of 1, 406 aa, 1 stop

<MW: 46927, pI: 5.21, NX(S/T): 0

MHPAVFLSLPDLRCSLLLLVTWVFTPVTTEITSLATENIDEILNNADVALVNFYADWCRFSQ
MLHPIFEEASDVIKEEFPNENQVVFARVDCDQHSDIAQRYRISKYPTLKLFRNGMMMKREYR
GQRSVKALADYIRQQKSDPIQEIRDLAEITTLDRSKRNIIGYFEQKDSDNYRVFERVANILH
DDCAFLSAFGDVSKPERYSGDNIIYKPPGHSAPDMVYLGAMTNFDVTYNWIQDKCVPLVREI
TFENGEELTEEGLPFLILFHMKEDTESLEIFQNEVARQLISEKGTINFLHADCDKFRHPLLH
IQKTPADCPVIAIDSFRHMYVFGDFKDVLIPGKLKQFVFDLHSGKLHREFHHGPDPTDTAPG
EQAQDVASSPPESSFQKLAPSEYRYTLLRDRDEL

Important features:

Signal peptide:

amino acids 1-29

Endoplasmic reticulum targeting sequence.

amino acids 403-406

Tyrosine kinase phosphorylation site.

amino acids 203-211

Thioredoxin family proteins

amino acids 50-66

FIGURE 126

ATTAAGGAAGAATTTCCAAATGAAAATCAAGTAGTNTTTGCCAGAGTNGATTGTGATCAGCA
CTCTGACATAGCCCAGAGATACAGGATAAGCAAATACCCAACCCTCAAATTGTTTCGTAATG
GGATGATGATGAAGAGAGAATACAGGGGTCAGCGATCAGTGAAAGCATTGGCAGATTA

FIGURE 127

AGAGGCCTCTCTGGAAGTTGTCCCGGGTGTTCGCCGCNGGAGCCCGGGTCGAGAGGACNAGG
TGCCGCTGCCTGGAGAATCCTCCGCTGCCGTCGGCTCCCGGAGCCCAGCCCTTTCCTAACCC
AACCCAACCTAGCCCNGTCCCAGCCGCCAGCGCCTGTCCCTGTCNCGGANCCCAGCGTNACC
ATGCATCCTGCCGTCTTCCTATCCTTACCCGACCTCAGATGCTCCCTTCTGCTCCTGGTAAC
TTGGGTTTTTACTCCTGTAACAACTGAAATAACNNGTCTTGATACNNAGAATATAGATGAAA
TTTTAAACNATGCTGATGTGGCTTTAGTCAATTTTTATGCTGACTGGTGTCGTTTCAGTCAG
ATGTGGCATCCAATTTTTGAGGANGCTTCCGATGTCATTAAGGAAGAATTTCCAAATGAAAA
TCAAGTAGTGTTTGCCAGAGTTGATTGTGATCAGCACTCTGACATAGCCCAGAGATACAGGA
TAAGCAAATACCCAACCCTCAAATTGTTTCGTAATGGGATGATGATGAAGAGAGAATACAGG
GGTCAGCGATCAGTGAAAGCATTGGCAGATTACATCAGGC

FIGURE 128

GCCCACGCGTCCGATGGCGTTCACGTTCGCGGCCTTCTGCTACATGCTGGCGCTGCTGCTCA
CTGCCGCGCTCATCTTCTTCGCCATTTGGCACATTATAGCATTTGATGAGCTGAAGACTGAT
TACAAGAATCCTATAGACCAGTGTAATACCCTGAATCCCCTTGTACTCCCAGAGTACCTCAT
CCACGCTTTCTTCTGTGTCATGTTTCTTTGTGCAGCAGAGTGGCTTACACTGGGTCTCAATA
TGCCCCTCTTGGCATATCATATTTGGAGGTATATGAGTAGACCAGTGATGAGTGGCCCAGGA
CTCTATGACCCTACAACCATCATGAATGCAGATATTCTAGCATATTGTCAGAAGGAAGGATG
GTGCAAATTAGCTTTTTATCTTCTAGCATTTTTTTACTACCTATATGGCATGATCTATGTTT
TGGTGAGCTCTTAGAACAACACACAGAAGAATTGGTCCAGTTAAGTGCATGCAAAAAGCCAC
CAAATGAAGGGATTCTATCCAGCAAGATCCTGTCCAAGAGTAGCCTGTGGAATCTGATCAGT
TACTTTAAAAAATGACTCCTTATTTTTTAAATGTTTCCACATTTTTGCTTGTGGAAAGACTG
TTTTCATATGTTATACTCAGATAAAGATTTTAAATGGTATTACGTATAAATTAATATAAAAT
GATTACCTCTGGTGTTGACAGGTTTGAACTTGCACTTCTTAAGGAACAGCCATAATCCTCTG
AATGATGCATTAATTACTGACTGTCCTAGTACATTGGAAGCTTTTGTTTATAGGAACTTGTA
GGGCTCATTTTGGTTTCATTGAAACAGTATCTAATTATAAATTAGCTGTAGATATCAGGTGC
TTCTGATGAAGTGAAAATGTATATCTGACTAGTGGGAAACTTCATGGGTTTCCTCATCTGTC
ATGTCGATGATTATATATGGATACATTTACAAAAATAAAAAGCGGGAATTTTCCCTTCGCTT
GAATATTATCCCTGTATATTGCATGAATGAGAGATTTCCCATATTTCCATCAGAGTAATAAA
TATACTTGCTTTAATTCTTAAGCATAAGTAAACATGATATAAAAATATATGCTGAATTACTT
GTGAAGAATGCATTTAAAGCTATTTTAAATGTGTTTTTATTTGTAAGACATTACTTATTAAG
AAATTGGTTATTATGCTTACTGTTCTAATCTGGTGGTAAAGGTATTCTTAAGAATTTGCAGG
TACTACAGATTTTCAAAACTGAATGAGAGAAAATTGTATAACCATCCTGCTGTTCCTTTAGT
GCAATACAATAAAACTCTGAAATTAAGACTC

FIGURE 129

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA23330

<subunit 1 of 1, 144 aa, 1 stop

<MW: 16699, pI: 5.60, NX(S/T): 0

MAFTFAAFCYMLALLLTAALIFFAIWHIIAFDELKTDYKNPIDQCNTLNPLVLPEYLIHAFF
CVMFLCAAEWLTLGLNMPLLAYHIWRYMSRPVMSGPGLYDPTTIMNADILAYCQKEGWCKLA
FYLLAFFYYLYGMIYVLVSS

Important features:

Signal peptide:

amino acids 1-20

Type II transmembrane domain:

amino acids 11-31

Other transmembrane domain:

amino acids 57-77 and 123-143

FIGURE 130

ATTATAGCATTTGATGAGCTGAAGACTGATTACAAGATCCTATAGACCAGTGTAATACCCTG
AATCCCCTTGTACTCCCAGAGTACCTCATCCACGCTTTCTTCTGTGTCATGTTTCTTTGTGC
AGCAGAGTGGCTTACACTGGGTCTCAATATGCCCCTCTTGGCATATCATATTTGGAGGTATA
TGAGTAGACCAGTGATGAGTGGCCCAGGACTCTATGACCCTACAACCATCATGAATGCAGAT
ATTCTAGCATATTGTCAGAAGGAAGGATGGTGCAAATTAGCTTTTTATCTTCTAGCATTTTT
TTACTACCTATATGGCATGATCTATGTTTTGGTGAGCTCTTAGAACAACACACAGAAGAATT
GGTCCAGTTAAGTGCATGCAAAAAGCCACCAAATGAAGGGATTCTATCCAGCAAGATCCTGT
CCAAGAGTAGCCTGTGGAATCTGATCAGTTACTTTAAAAAATG

FIGURE 131

CGGACGCGTGGGGGAAACCCTTCCGAGAAAACAGCAACAAGCTGAGCTGCTGTGACAGAGGG
GAACAAGATGGCGGCGCCGAAGGGGAGCCTCTGGGTGAGGACCCAACTGGGGCTCCCGCCGC
TGCTGCTGCTGACCATGGCCTTGGCCGGAGGTTCGGGGACCGCTTCGGCTGAAGCATTTGAC
TCGGTCTTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAGTTGACCTACCCCTTGCACAC
CTACCCTAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTGCAGGCTGTTTTCAATTTGTC
AGTTTGTGGATGATGGAATTGACTTAAATCGAACTAAATTGGAATGTGAATCTGCATGTACA
GAAGCATATTCCCAATCTGATGAGCAATATGCTTGCCATCTTGGTTGCCAGAATCAGCTGCC
ATTCGCTGAACTGAGACAAGAACAACTTATGTCCCTGATGCCAAAAATGCACCTACTCTTTC
CTCTAACTCTGGTGAGGTCATTCTGGAGTGACATGATGGACTCCGCACAGAGCTTCATAACC
TCTTCATGGACTTTTTATCTTCAAGCCGATGACGGAAAAATAGTTATATTCCAGTCTAAGCC
AGAAATCCAGTACGCACCACATTTGGAGCAGGAGCCTACAAATTTGAGAGAATCATCTCTAA
GCAAAATGTCCTATCTGCAAATGAGAAATTCACAAGCGCACAGGAATTTTCTTGAAGATGGA
GAAAGTGATGGCTTTTTAAGATGCCTCTCTCTTAACTCTGGGTGGATTTTAACTACAACTCT
TGTCCTCTCGGTGATGGTATTGCTTTGGATTTGTTGTGCAACTGTTGCTACAGCTGTGGAGC
AGTATGTTCCCTCTGAGAAGCTGAGTATCTATGGTGACTTGGAGTTTATGAATGAACAAAAG
CTAAACAGATATCCAGCTTCTTCTCTTGTGGTTGTTAGATCTAAAACTGAAGATCATGAAGA
AGCAGGGCCTCTACCTACAAAAGTGAATCTTGCTCATTCTGAAATTTAAGCATTTTTCTTTT
AAAAGACAAGTGTAATAGACATCTAAAATTCCACTCCTCATAGAGCTTTTAAAATGGTTTCA
TTGGATATAGGCCTTAAGAAATCACTATAAAATGCAAATAAAGTTACTCAAATCTGTG

FIGURE 132

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA26847

<subunit 1 of 1, 323 aa, 1 stop

<MW: 36223, pI: 5.06, NX(S/T): 1

MAAPKGSLWVRTQLGLPPLLLLTMALAGGSGTASAEAFDSVLGDTASCHRACQLTYPLHTYP
KEEELYACQRGCRLFSICQFVDDGIDLNRTKLECESACTEAYSQSDEQYACHLGCQNQLPFA
ELRQEQLMSLMPKMHLLFPLTLVRSFWSDMMDSAQSFITSSWTFYLQADDGKIVIFQSKPEI
QYAPHLEQEPTNLRESSLSKMSYLQMRNSQAHRNFLEDGESDGFLRCLSLNSGWILTTTLVL
SVMVLLWICCATVATAVEQYVPSEKLSIYGDLEFMNEQKLNRYPASSLVVVRSKTEDHEEAG
PLPTKVNLAHSEI

Important features:

Signal peptide:

amino acids 1-31

Transmembrane domain:

amino acids 241-260

N-glycosylation site.

amino acids 90-93

FIGURE 133

TTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAGTTGACCTACCCCTTGCACACCTACCC
TAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTGCAGGCTGTTTTCAATTTGTCAGTTTG
TGGATGATGGAATTGACTTAAATCGAACTAAATTGGAATGTGAATCTGCATGTACAGAAGCA
TATTCCCAATCTGATGAGCAATATGCTTGCCATCTTGGTTGCCAGAATCAGCTGCCATTCGC
TGAACTGAGACAAGAACAACTTATGTCCCTGATGCCAAAAATGCACCTACTCTTTCCTCTAA
CTCTGGTGAGGTCATTCTGGAGTGACATGATGGACTCCGC

FIGURE 134

CACACTGGCCGGATCTTTTAGAGTCCTTTGACCTTGACCAAGGGTCNGGAAAACAGCAACAA
GCTGAGCTGCTGTGACAGAGGGAACAAGATGGCGGCGCCGAAGGGAGCCTTTGGGTGAGGAC
CCAACTGGGGCTCCCGCCGCTGCTGCTGCTGACCATGGCCTTGGCCGGAGGTTCGGGGACCG
CTTCGGCTGAAGCATTTGACTCGGTCTTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAG
TTGACCTACCCCTTGCACACCTACCCTAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTG
CAGGCTGTTTTCAATTTGTCAGTTTGTGGATGATGGAATTGACTTAAATCGAACTAAATTGG
AATGTGAATCTGCATGTACAGAAGCATATTCCCAATCTGATGAGCAATATGCTTGCCATCTT
GGTTGCCAGAATCAGCTGCCATTCGCTGAACTGAGACAAGAACAACTTATGTCCCTGATGCC
AAAAATGCACCTACTCTTTCCTCTAACTCTGGTGAGGTCATTCTGGAGTGACATGATGGACT
CCGC

FIGURE 135

GCGAGGTGGCGATCGCTGAGAGGCAGGAGGGCCGAGGCGGGCCTGGGAGGCGGCCCGGAGGT
GGGGCGCCGCTGGGGCCGGCCCGCACGGGCTTCATCTGAGGGCGCACGGCCCGCGACCGAGC
GTGCGGACTGGCCTCCCAAGCGTGGGGCGACAAGCTGCCGGAGCTGCAATGGGCCGCGGCTG
GGGATTCTTGTTTGGCCTCCTGGGCGCCGTGTGGCTGCTCAGCTCGGGCCACGGAGAGGAGC
AGCCCCCGGAGACAGCGGCACAGAGGTGCTTCTGCCAGGTTAGTGGTTACTTGGATGATTGT
ACCTGTGATGTTGAAACCATTGATAGATTTAATAACTACAGGCTTTTCCCAAGACTACAAAA
ACTTCTTGAAAGTGACTACTTTAGGTATTACAAGGTAAACCTGAAGAGGCCGTGTCCTTTCT
GGAATGACATCAGCCAGTGTGGAAGAAGGGACTGTGCTGTCAAACCATGTCAATCTGATGAA
GTTCCTGATGGAATTAAATCTGCGAGCTACAAGTATTCTGAAGAAGCCAATAATCTCATTGA
AGAATGTGAACAAGCTGAACGACTTGGAGCAGTGGATGAATCTCTGAGTGAGGAAACACAGA
AGGCTGTTCTTCAGTGGACCAAGCATGATGATTCTTCAGATAACTTCTGTGAAGCTGATGAC
ATTCAGTCCCCTGAAGCTGAATATGTAGATTTGCTTCTTAATCCTGAGCGCTACACTGGTTA
CAAGGGACCAGATGCTTGGAAAATATGGAATGTCATCTACGAAGAAAACTGTTTTAAGCCAC
AGACAATTAAAAGACCTTTAAATCCTTTGGCTTCTGGTCAAGGGACAAGTGAAGAGAACACT
TTTTACAGTTGGCTAGAAGGTCTCTGTGTAGAAAAAAGAGCATTCTACAGACTTATATCTGG
CCTACATGCAAGCATTAATGTGCATTTGAGTGCAAGATATCTTTTACAAGAGACCTGGTTAG
AAAAGAAATGGGGACACAACATTACAGAATTTCAACAGCGATTTGATGGAATTTTGACTGAA
GGAGAAGGTCCAAGAAGGCTTAAGAACTTGTATTTTCTCTACTTAATAGAACTAAGGGCTTT
ATCCAAAGTGTTACCATTCTTCGAGCGCCCAGATTTTCAACTCTTTACTGGAAATAAAATTC
AGGATGAGGAAAACAAAATGTTACTTCTGGAAATACTTCATGAAATCAAGTCATTTCCTTTG
CATTTTGATGAGAATTCATTTTTTGCTGGGGATAAAAAAGAAGCACACAAACTAAAGGAGGA
CTTTCGACTGCATTTTAGAAATATTTCAAGAATTATGGATTGTGTTGGTTGTTTTAAATGTC
GTCTGTGGGGAAAGCTTCAGACTCAGGGTTTGGGCACTGCTCTGAAGATCTTATTTTCTGAG
AAATTGATAGCAAATATGCCAGAAAGTGGACCTAGTTATGAATTCCATCTAACCAGACAAGA
AATAGTATCATTATTCAACGCATTTGGAAGAATTTCTACAAGTGTGAAAGAATTAGAAAACT
TCAGGAACTTGTTACAGAATATTCATTAAAGAAAACAAGCTGATATGTGCCTGTTTCTGGAC
AATGGAGGCGAAAGAGTGGAATTTCATTCAAAGGCATAATAGCAATGACAGTCTTAAGCCAA
ACATTTTATATAAAGTTGCTTTTGTAAAGGAGAATTATATTGTTTTAAGTAAACACATTTTT
AAAAATTGTGTTAAGTCTATGTATAATACTACTGTGAGTAAAAGTAATACTTTAATAATGTG
GTACAAATTTTAAAGTTTAATATTGAATAAAAGGAGGATTATCAAATTAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 136

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53974

<subunit 1 of 1, 468 aa, 1 stop

<MW: 54393, pI: 5.63, NX(S/T): 2

MGRGWGFLFGLLGAVWLLSSGHGEEQPPETAAQRCFCQVSGYLDDCTCDVETIDRFNNYRLF
PRLQKLLESDYFRYYKVNLKRPCPFWNDISQCGRRDCAVKPCQSDEVPDGIKSASYKYSEEA
NNLIEECEQAERLGAVDESLSEETQKAVLQWTKHDDSSDNFCEADDIQSPEAEYVDLLLNPE
RYTGYKGPDAWKIWNVIYEENCFKPQTIKRPLNPLASGQGTSEENTFYSWLEGLCVEKRAFY
RLISGLHASINVHLSARYLLQETWLEKKWGHNITEFQQRFDGILTEGEGPRRLKNLYFLYLI
ELRALSKVLPFFERPDFQLFTGNKIQDEENKMLLLEILHEIKSFPLHFDENSFFAGDKKEAH
KLKEDFRLHFRNISRIMDCVGCFKCRLWGKLQTQGLGTALKILFSEKLIANMPESGPSYEFH
LTRQEIVSLFNAFGRISTSVKELENFRNLLQNIH

Important features:

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 280-283 and 384-387

Amidation site.

amino acids 94-97

Glycosaminoglycan attachment site.

amino acids 20-23 and 223-226

Aminotransferases class-V pyridoxal-phosphate

amino acids 216-222

Interleukin-7 proteins

amino acids 338-343

FIGURE 137

GCTGGAAATATGGATGTCATCTACGAGAAACTGTTTTAAGCCACAGACAATTAAAAGACCTT
TAAATCCTTTGGCTTCTGGTCAAGGGACAAGTGAAGAGNACACTTTTTACAGTTGGCTAGAA
GGTCTCTGTGTAGAAAAAAGAGCATTCTACAGACTTATATCTGGCCTACATGCAAGCATTAA
TGTGCATTTGAGTGCAAGATATCTTTTACAAGAGACCTGGTTAGAAAAGAAATGGGGACACA
ACATTACAGAATTTNAACAGCGATTTGATGGAATTTTGACTGAAGGAGAAGGTCCAAGAAGG
CTTAAGAACTTGTATTTTCTCTACTTAATAGAACTAAGGGCTTTATCCAAAGTGTTACCATT
CTTNGAGCGCCCAGATTTTCAACTNTTTACTGGAAATAAAATTCAGGATGAGGNAAACAAAA
TGTTACTTTTGGAAATACTTCATGAAATCAAGTCATTTCCTTTGCATTTTGATGAGAATTCA
TTTTTTTGCTG

FIGURE 138

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGTTGGGAGGGGGCAGGATGGGAGGGAA
AGTGAAGAAAACAGAAAAGGAGAGGGACAGAGGCCAGAGGACTTCTCATACTGGACAGAAAC
CGATCAGGCATGGAACTCCCCTTCGTCACTCACCTGTTCTTGCCCCTGGTGTTCCTGACAGG
TCTCTGCTCCCCCTTTAACCTGGATGAACATCACCCACGCCTATTCCCAGGGCCACCAGAAG
CTGAATTTGGATACAGTGTCTTACAACATGTTGGGGGTGGACAGCGATGGATGCTGGTGGGC
GCCCCCTGGGATGGGCCTTCAGGCGACCGGAGGGGGGACGTTTATCGCTGCCCTGTAGGGGG
GGCCCACAATGCCCCATGTGCCAAGGGCCACTTAGGTGACTACCAACTGGGAAATTCATCTC
ATCCTGCTGTGAATATGCACCTGGGGATGTCTCTGTTAGAGACAGATGGTGATGGGGGATTC
ATGGTGAGCTAAGGAGAGGGTGGTGGCAGTGTCTCTGAAGGTCCATAAAAGAAAAAAGAGAA
GTGTGGTAAGGGAAAATGGTCTGTGTGGAGGGGTCAAGGAGTTAAAAACCCTAGAAAGCAAA
AGGTAGGTAATGTCAGGGAGTAGTCTTCATGCCTCCTTCAACTGGGAGCATGTTCTGAGGGT
GCCCTCCCAAGCCTGGGAGTAACTATTTCCCCCATCCCCAGGCCTGTGCCCCTCTCTGGTCT
CGTGCTTGTGGCAGCTCTGTCTTCAGTTCTGGGATATGTGCCCGTGTGGATGCTTCATTCCA
GCCTCAGGGAAGCCTGGCACCCACTGCCCAACGTGAGCCAGAGGAAGGCTGAGTACTTGGTT
CCCAGAAGGAGATACTGGGTGGGAAAAAGATGGGGCAAAGCGGTATGATGCCTGGCAAAGGG
CCTGCATGGCTATCCTCATTGCTACCTAATGTGCTTGCAAAAGCTCCATGTTTCCTAACAGA
TTCAGACTCCTGGCCAGGTGTGGTGGCCCACACCTGTAATTCTAGCACTTTGGGAGGCCAAG
GTGGGCAGATCACTTGAGGTCAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACTCCAT
CTCTACTAAAAAAAAAAAAATACAAAAATTAGCTGGGTGCGCTAGTGCATGCCTGTAATCTC
ATCTACTCGGGAGGCTAAGACAGGAGACTCTCACTTCAACCCAGGAGGTGGAGGTTGCGGTG
AGCCAAGATTGTGCCTCTGCACTCTAGCGTGGGTGACAGAGTAAGCGAGACTCCATCTCAAA
AATAATAATAATAATAATTCAGACTCCTTATCAGGAGTCCATGATCTGGCCTGGCACAGTAA
CTCATGCCTGTAATCCCAACATTTTGGGAGGCCAACGCAGGAGGATTGCTTGAGGTCTGGAG
GTTTGAGACCAGCCTGGGCAACATAGAAAGACCCCATCTCTAAATAAATGTTTTAAAAAT

FIGURE 139

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57039

><subunit 1 of 1, 124 aa, 1 stop

><MW: 13352, pI: 5.99, NX(S/T): 1

MELPFVTHLFLPLVFLTGLCSPFNLDEHHPRLFPGPPEAEFGYSVLQHVGGGQRWMLVGAPW

DGPSGDRRGDVYRCPVGGAHNAPCAKGHLGDYQLGNSSHPAVNMHLGMSLLETDGDGGFMVS

Important features:

Signal peptide:

amino acids 1-22

Cell attachment sequence.

amino acids 70-73

N-glycosylation site.

amino acids 98-101

Integrins alpha chain proteins

amino acids 67-81

FIGURE 140

CACAGTTCCCCACCATCACTCNTCCCATTCCTTCCAACTTTATTTTTAGCTTGCCATTGGGA
GGGGGCAGGATGGGAGGGAAAGTGAAGAAAACAGAAAAGGAGAGGGACAGAGGCCAGAGGAC
TTCTCATACTGGACAGAAACCGATCAGGCATGGAACTCCCCTTCGTCACTCACCTGTTCTTG
CCCCTGGTGTTCCTGACAGGTCTCTGCTCCCCCTTTAACCTGGATGAACATCACCCACGCCT
ATTCCCAGGGCCACCAGAAGCTGAATTTGGATACAGTGTCTTACAACATGTTGGGGGTGGAC
AGCGATGGATGCTGGTGGGCGCCCCCTGGGATGGGCCTTCAGGCGACCGGAGGGGGGACGTT
TATCGCTGCCCTGTAGGGGGGGCCCACAATGCCCCATGTGCCAAGGGCCACTTAGGTGACTA
CCAACTGGGAAATTCATCTCATCCTGCTGTGAATATGCACCTGGGGATGTCTCTGTTAGAGA
CAGATGGTGATGG

FIGURE 141

AAAGTTACATTTTCTCTGGAACTCTCCTAGGCCACTCCCTGCTGATGCAACATCTGGGTTTG
GGCAGAAAGGAGGGTGCTTCGGAGCCCGCCCTTTCTGAGCTTCCTGGGCCGGCTCTAGAACA
ATTCAGGCTTCGCTGCGACTCAGACCTCAGCTCCAACATATGCATTCTGAAGAAAGATGGCT
GAGATGGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGGTCAAACTGAGTCTACCA
AATGCAGACTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCTTTTCATGTGGTTTTTCT
ACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGCCTGCCCCTCAGAACCTC
TCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCAGTGATCGCGCCTGGAGA
AACAGTGTACTATTCTGTCGAATACCAGGGGGAGTACGAGAGCCTGTACACGAGCCACATCT
GGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCCTGAGTGTGATGTCACTGATGACATC
ACGGCCACTGTGCCATACAACCTTCGTGTCAGGGCCACATTGGGCTCACAGACCTCAGCCTG
GAGCATCCTGAAGCATCCCTTTAATAGAAACTCAACCATCCTTACCCGACCTGGGATGGAGA
TCACCAAAGATGGCTTCCACCTGGTTATTGAGCTGGAGGACCTGGGGCCCCAGTTTGAGTTC
CTTGTGGCCTACTGGAGGAGGGAGCCTGGTGCCGAGGAACATGTCAAAATGGTGAGGAGTGG
GGGTATTCCAGTGCACCTAGAAACCATGGAGCCAGGGGCTGCATACTGTGTGAAGGCCCAGA
CATTCGTGAAGGCCATTGGGAGGTACAGCGCCTTCAGCCAGACAGAATGTGTGGAGGTGCAA
GGAGAGGCCATTCCCCTGGTACTGGCCCTGTTTGCCTTTGTTGGCTTCATGCTGATCCTTGT
GGTCGTGCCACTGTTCGTCTGGAAAATGGGCCGGCTGCTCCAGTACTCCTGTTGCCCCGTGG
TGGTCCTCCCAGACACCTTGAAAATAACCAATTCACCCCAGAAGTTAATCAGCTGCAGAAGG
GAGGAGGTGGATGCCTGTGCCACGGCTGTGATGTCTCCTGAGGAACTCCTCAGGGCCTGGAT
CTCATAGGTTTGCGGAAGGGCCCAGGTGAAGCCGAGAACCTGGTCTGCATGACATGGAAACC
ATGAGGGGACAAGTTGTGTTTCTGTTTTCCGCCACGGACAAGGGATGAGAGAAGTAGGAAGA
GCCTGTTGTCTACAAGTCTAGAAGCAACCATCAGAGGCAGGGTGGTTTGTCTAACAGAACAC
TGACTGAGGCTTAGGGGATGTGACCTCTAGACTGGGGGCTGCCACTTGCTGGCTGAGCAACC
CTGGGAAAAGTGACTTCATCCCTTCGGTCCTAAGTTTTCTCATCTGTAATGGGGGAATTACC
TACACACCTGCTAAACACACACACACAGAGTCTCTCTCTATATATACACACGTACACATAAA
TACACCCAGCACTTGCAAGGCTAGAGGGAAACTGGTGACACTCTACAGTCTGACTGATTCAG
TGTTTCTGGAGAGCAGGACATAAATGTATGATGAGAATGATCAAGGACTCTACACACTGGGT
GGCTTGGAGAGCCCACTTTCCCAGAATAATCCTTGAGAGAAAAGGAATCATGGGAGCAATGG
TGTTGAGTTCACTTCAAGCCCAATGCCGGTGCAGAGGGGAATGGCTTAGCGAGCTCTACAGT
AGGTGACCTGGAGGAAGGTCACAGCCACACTGAAAATGGGATGTGCATGAACACGGAGGATC
CATGAACTACTGTAAAGTGTTGACAGTGTGTGCACACTGCAGACAGCAGGTGAAATGTATGT
GTGCAATGCGACGAGAATGCAGAAGTCAGTAACATGTGCATGTTTGTTGTGCTCCTTTTTTC
TGTTGGTAAAGTACAGAATTCAGCAAATAAAAAGGGCCACCCTGGCCAAAAGCGGTAAAAAA
AAAAAAAAAA

FIGURE 142

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57033
<subunit 1 of 1, 311 aa, 1 stop
<MW: 35076, pI: 5.04, NX(S/T): 2
MQTFTMVLEEIWTSLFMWFFYALIPCLLTDEVAILPAPQNLSVLSTNMKHLLMWSPVIAPGE
TVYYSVEYQGEYESLYTSHIWIPSSWCSLTEGPECDVTDDITATVPYNLRVRATLGSQTSAW
SILKHPFNRNSTILTRPGMEITKDGFHLVIELEDLGPQFEFLVAYWRREPGAEEHVKMVRSG
GIPVHLETMEPGAAYCVKAQTFVKAIGRYSAFSQTECVEVQGEAIPLVLALFAFVGFMLILV
VVPLFVWKMGRLLQYSCCPVVVLPDTLKITNSPQKLISCRREEVDACATAVMSPEELLRAWIS

Important features:

Signal peptide:
amino acids 1-29

Transmembrane domain:
amino acids 230-255

N-glycosylation site.
amino acids 40-43 and 134-137

Tissue factor proteins.
amino acids 92-119

Integrins alpha chain proteins
amino acids 232-262

FIGURE 143

TCCTGCTGATGCACATCTGGGTTTGGCAAAAGGAGGTTGCTTCGAGCCGCCCTTTCTAGCTT
CCTGGCCGGCTCTAGAACAATTCAGGCTTCGCTGCGACTAGACCTCAGCTCCAACATATGCA
TTCTGAAGAAAGATGGCTGAGATGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGG
TCAAACTGAGTCTACCAAATGCAGACTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCT
TTTCATGTGGTTTTTCTACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGC
CTGCCCCTCAGAACCTCTCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCA
GTGATCGCGCCTGGAGAAACAGTGTACTATTCTGTCGAATACCAGGGGGAGTACGAGAGCCT
GTACACGAGCCACATCTGGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCCTGAGTGTG
ATGTCACTGATGACATCACGGCCACTGTGCCATACAACCTTTGTGTCAGGGCCACATTGGGC
TCACAGACCTCAGCCTGGAGCATCCTGAAGCATCCCTTTAATAGAAACTCAACCATCCTTAC
CCGACCTGGGATGGAGATCACCAAAGATGGCTTNCACCTGGTTATTGAGCTGGAGGACCTGG
GGCCCCAGTTTGAGTTCCTTGTGGCCTANTGGAGGAGGGGCGAACCCCTTGCGGCGCAAGGG
GTTNGCGAACCCCTTGCGGCCGCTGGGGTATCTCTCGAGAAAAGAGAGGCCCAATATGACCC
ACATACTCAATATGGACGAANTGCTATTGTCCACCTGTTTGAGTGGCGCTGGGTTGAT

FIGURE 144

CCCACGCGTCCGCCCACGCGTCCGAGGGACAAGAGAGAAGAGAGACTGAAACAGGGAGAAGA
GGCAGGAGAGGAGGAGGTGGGGAGAGCACGAAGCTGGAGGCCGACACTGAGGGAGGGCGGGA
GGAGGTGAAGAAGGAGAGAGGGGAGAAGAGGCAGGAGCTGGAAAGGAGAGAGGGAGGAGGAG
GAGGAGATGCGGGATGGAGACCTGGAGTTAGGTGGCTTGGGAGAGCTTAATGAAAAGAGAAC
GGAGAGGAGGTGTGGGTTAGGAACCAAGAGGTAGCCCTGTGGGCAGCAGAAGGCTGAGAGGA
GTAGGAAGATCAGGAGCTAGAGGGAGACTGGAGGGTTCCGGGAAAAGAGCAGAGGAAAGAGG
AAAGACACAGAGAGACGGGAGAGAGAAGAAGAGTGGGTTTGAAGGGCGGATCTCAGTCCCTG
GCTGCTTTGGCATTTGGGGAACTGGGACTCCCTGTGGGGAGGAGAGGAAAGCTGGAAGTCCT
GGAGGGACAGGGTCCCAGAAGGAGGGGACAGAGGAGCTGAGAGAGGGGGGCAGGGCGTTGGG
CAGGGGTCCCTCGGAGGCCTCCTGGGG**ATG**GGGGCTGCAGCTCGTCTGAGCGCCCCTCGAGC
GCTGGTACTCTGGGCTGCACTGGGGGCAGCAGCTCACATCGGACCAGCACCTGACCCCGAGG
ACTGGTGGAGCTACAAGGATAATCTCCAGGGAAACTTCGTGCCAGGGCCTCCTTTCTGGGGC
CTGGTGAATGCAGCGTGGAGTCTGTGTGCTGTGGGGAAGCGGCAGAGCCCCGTGGATGTGGA
GCTGAAGAGGGTTCTTTATGACCCCTTTCTGCCCCCATTAAGGCTCAGCACTGGAGGAGAGA
AGCTCCGGGGAACCTTGTACAACACCGGCCGACATGTCTCCTTCCTGCCTGCACCCCGACCT
GTGGTCAATGTGTCTGGAGGTCCCCTCCTTTACAGCCACCGACTCAGTGAACTGCGGCTGCT
GTTTGGAGCTCGCGACGGAGCCGGCTCGGAACATCAGATCAACCACCAGGGCTTCTCTGCTG
AGGTGCAGCTCATTCACTTCAACCAGGAACTCTACGGGAATTTCAGCGCTGCCTCCCGCGGC
CCCAATGGCCTGGCCATTCTCAGCCTCTTTGTCAACGTTGCCAGTACCTCTAACCCATTCCT
CAGTCGCCTCCTTAACCGCGACACCATCACTCGCATCTCCTACAAGAATGATGCCTACTTTC
TTCAAGACCTGAGCCTGGAGCTCCTGTTCCCTGAATCCTTCGGCTTCATCACCTATCAGGGC
TCTCTCAGCACCCCGCCCTGCTCCGAGACTGTCACCTGGATCCTCATTGACCGGGCCCTCAA
TATCACCTCCCTTCAGATGCACTCCCTGAGACTCCTGAGCCAGAATCCTCCATCTCAGATCT
TCCAGAGCCTCAGCGGTAACAGCCGGCCCCTGCAGCCCTTGGCCCACAGGGCACTGAGGGGC
AACAGGGACCCCCGGCACCCCGAGAGGCGCTGCCGAGGCCCCAACTACCGCCTGCATGTGGA
TGGTGTCCCCCATGGTCGC**TGA**GACTCCCCTTCGAGGATTGCACCCGCCCGTCCTAAGCCTC
CCCACAAGGCGAGGGGAGTTACCCCTAAAACAAAGCTATTAAAGGGACAGAATACTTA

FIGURE 145

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA34353
<subunit 1 of 1, 328 aa, 1 stop
<MW: 36238, pI: 9.90, NX(S/T): 3
MGAAARLSAPRALVLWAALGAAAHIGPAPDPEDWWSYKDNLQGNFVPGPPFWGLVNAAWSLC
AVGKRQSPVDVELKRVLYDPFLPPLRLSTGGEKLRGTLYNTGRHVSFLPAPRPVVNVSGGPL
LYSHRLSELRLLFGARDGAGSEHQINHQGFSAEVQLIHFNQELYGNFSAASRGPNGLAILSL
FVNVASTSNPFLSRLLNRDTITRISYKNDAYFLQDLSLELLFPESFGFITYQGSLSTPPCSE
TVTWILIDRALNITSLQMHSLRLLSQNPPSQIFQSLSGNSRPLQPLAHRALRGNRDPRHPER
RCRGPNYRLHVDGVPHGR

Important features:

Signal peptide:
amino acids 1-23

Transmembrane domain:
amino acids 177-199

N-glycosylation site.
amino acids 118-121, 170-173 and 260-263

Eukaryotic-type carbonic anhydrases proteins
amino acids 222-270, 128-164 and 45-92

FIGURE 146

GGCGCCTGGTTCTGCGCGTACTGGCTGTACGGAGCAGGAGCAAGAGGTCGCCGCCAGCCTCC
GCCGCCGAGCCTCGTTCGTGTCCCCGCCCCTCGCTCCTGCAGCTACTGCTCAGAAACGCTGG
GGCGCCCACCCTGGCAGACTAACGAAGCAGCTCCCTTCCCACCCCAACTGCAGGTCTAATTT
TGGACGCTTTGCCTGCCATTTCTTCCAGGTTGAGGGAGCCGCAGAGGCGGAGGCTCGCGTAT
TCCTGCAGTCAGCACCCACGTCGCCCCCGGACGCTCGGTGCTCAGGCCCTTCGCGAGCGGGG
CTCTCCGTCTGCGGTCCCTTGTGAAGGCTCTGGGCGGCTGCAGAGGCCGGCCGTCCGGTTTG
GCTCACCTCTCCCAGGAAACTTCACACTGGAGAGCCAAAAGGAGTGGAAGAGCCTGTCTTGG
AGATTTTCCTGGGGAAATCCTGAGGTCATTCATTATGAAGTGTACCGCGCGGGAGTGGCTCA
GAGTAACCACAGTGCTGTTCATGGCTAGAGCAATTCCAGCCATGGTGGTTCCCAATGCCACT
TTATTGGAGAAACTTTTGGAAAAATACATGGATGAGGATGGTGAGTGGTGGATAGCCAAACA
ACGAGGGAAAAGGGCCATCACAGACAATGACATGCAGAGTATTTTGGACCTTCATAATAAAT
TACGAAGTCAGGTGTATCCAACAGCCTCTAATATGGAGTATATGACATGGGATGTAGAGCTG
GAAAGATCTGCAGAATCCTGGGCTGAAAGTTGCTTGTGGGAACATGGACCTGCAAGCTTGCT
TCCATCAATTGGACAGAATTTGGGAGCACACTGGGGAAGATATAGGCCCCCGACGTTTCATG
TACAATCGTGGTATGATGAAGTGAAAGACTTTAGCTACCCATATGAACATGAATGCAACCCA
TATTGTCCATTCAGGTGTTCTGGCCCTGTATGTACACATTATACACAGGTCGTGTGGGCAAC
TAGTAACAGAATCGGTTGTGCCATTAATTTGTGTCATAACATGAACATCTGGGGGCAGATAT
GGCCCAAAGCTGTCTACCTGGTGTGCAATTACTCCCCAAAGGGAAACTGGTGGGGCCATGCC
CCTTACAAACATGGGCGGCCCTGTTCTGCTTGCCCACCTAGTTTTGGAGGGGGCTGTAGAGA
AAATCTGTGCTACAAAGAAGGGTCAGACAGGTATTATCCCCCTCGAGAAGAGGAAACAAATG
AAATAGAACGACAGCAGTCACAAGTCCATGACACCCATGTCCGGACAAGATCAGATGATAGT
AGCAGAAATGAAGTCATAAGCGCACAGCAAATGTCCCAAATTGTTTCTTGTGAAGTAAGATT
AAGAGATCAGTGCAAAGGAACAACCTGCAATAGGTACGAATGTCCTGCTGGCTGTTTGGATA
GTAAAGCTAAAGTTATTGGCAGTGTACATTATGAAATGCAATCCAGCATCTGTAGAGCTGCA
ATTCATTATGGTATAATAGACAATGATGGTGGCTGGGTAGATATCACTAGACAAGGAAGAAA
GCATTATTTCATCAAGTCCAATAGAAATGGTATTCAAACAATTGGCAAATATCAGTCTGCTA
ATTCCTTCACAGTCTCTAAAGTAACAGTTCAGGCTGTGACTTGTGAAACAACTGTGGAACAG
CTCTGTCCATTTCATAAGCCTGCTTCACATTGCCCAAGAGTATACTGTCCTCGTAACTGTAT
GCAAGCAAATCCACATTATGCTCGTGTAATTGGAACTCGAGTTTATTCTGATCTGTCCAGTA
TCTGCAGAGCAGCAGTACATGCTGGAGTGGTTCGAAATCACGGTGGTTATGTTGATGTAATG
CCTGTGGACAAAAGAAAGACCTACATTGCTTCTTTTCAGAATGGAATCTTCTCAGAAAGTTT
ACAGAATCCTCCAGGAGGAAAGGCATTCAGAGTGTTTGCTGTTGTGTGAAACTGAATACTTG
GAAGAGGACCATAAAGACTATTCCAAATGCAATATTTCTGAATTTTGTATAAAACTGTAACA
TTACTGTACAGAGTACATCAACTATTTTCAGCCCAAAAAGGTGCCAAATGCATATAAATCTT
GATAAACAAAGTCTATAAAATAAAACATGGGACATTAGCTTTGGGAAAAGTAATGAAAATAT
AATGGTTTTAGAAATCCTGTGTTAAATATTGCTATATTTTCTTAGCAGTTATTTCTACAGTT
AATTACATAGTCATGATTGTTCTACGTTTCATATATTATATGGTGCTTTGTATATGCCACTA
ATAAAATGAATCTAAACATTGAATGTGAATGGCCCTCAGAAAATCATCTAGTGCATTTAAAA
ATAATCGACTCTAAAACTGAAAGAAACCTTATCACATTTTCCCCAGTTCAATGCTATGCCAT
TACCAACTCCAAATAATCTCAAATAATTTTCCACTTAATAACTGTAAAGTTTTTTTCTGTTA
ATTTAGGCATATAGAATATTAAATTCTGATATTGCACTTCTTATTTTATATAAAATAATCCT
TTAATATCCAAATGAATCTGTTAAAATGTTTGATTCCTTGGGAATGGCCTTAAAAATAAATG
TAATAAAGTCAGAGTGGTGGTATGAAAACATTCCTAGTGATCATGTAGTAAATGTAGGGTTA
AGCATGGACAGCCAGAGCTTTCTATGTACTGTTAAAATTGAGGTCACATATTTTCTTTTGTA
TCCTGGCAAATACTCCTGCAGGCCAGGAAGTATAATAGCAAAAAGTTGAACAAAGATGAACT
AATGTATTACATTACCATTGCCACTGATTTTTTTTAAATGGTAAATGACCTTGTATATAAAT
ATTGCCATATCATGGTACCTATAATGGTGATATATTTGTTTCTATGAAAAATGTATTGTGCT
TTGATACTAAAAATCTGTAAAATGTTAGTTTTGGTAATTTTTTTTCTGCTGGTGGATTTACA
TATTAAATTTTTTCTGCTGGTGGATAAACATTAAAATTAATCATGTTTCAAAAAAAAAAAAA

FIGURE 147

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45417
<subunit 1 of 1, 500 aa, 1 stop
<MW: 56888, pI: 8.53, NX(S/T): 2
MKCTAREWLRVTTVLFMARAIPAMVVPNATLLEKLLEKYMDEDGEWWIAKQRGKRAITDNDM
QSILDLHNKLRSQVYPTASNMEYMTWDVELERSAESWAESCLWEHGPASLLPSIGQNLGAHW
GRYRPPTFHVQSWYDEVKDFSYPYEHECNPYCPFRCSGPVCTHYTQVVWATSNRIGCAINLC
HNMNIWGQIWPKAVYLVCNYSPKGNWWGHAPYKHGRPCSACPPSFGGGCRENLCYKEGSDRY
YPPREEETNEIERQQSQVHDTHVRTRSDDSSRNEVISAQQMSQIVSCEVRLRDQCKGTTCNR
YECPAGCLDSKAKVIGSVHYEMQSSICRAAIHYGIIDNDGGWVDITRQGRKHYFIKSNRNGI
QTIGKYQSANSFTVSKVTVQAVTCETTVEQLCPFHKPASHCPRVYCPRNCMQANPHYARVIG
TRVYSDLSSICRAAVHAGVVRNHGGYVDVMPVDKRKTYIASFQNGIFSESLQNPPGGKAFRV
FAVV

Important features:

Signal peptide:
amino acids 1-20

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 protein
amino acids 165-186, 196-218, 134-146, 96-108 and 58-77

N-glycosylation site
amino acids 28-31

FIGURE 148

GCGGAGACAAGCGCAGAGCGCAGCGCACGGCCACAGACAGCCCTGGGCATCCACCGACGGCG
CAGCCGGAGCCAGCAGAGCCGGAAGGCGCGCCCCGGGCAGAGAAAGCCGAGCAGAGCTGGGT
GGCGTCTCCGGGCCGCCGCTCCGACGGGCCAGCGCCCTCCCCATGTCCCTGCTCCCACGCCG
CGCCCCTCCGGTCAGCATGAGGCTCCTGGCGGCCGCGCTGCTCCTGCTGCTGCTGGCGCTGT
ACACCGCGCGTGTGGACGGGTCCAAATGCAAGTGCTCCCGGAAGGGACCCAAGATCCGCTAC
AGCGACGTGAAGAAGCTGGAAATGAAGCCAAAGTACCCGCACTGCGAGGAGAAGATGGTTAT
CATCACCACCAAGAGCGTGTCCAGGTACCGAGGTCAGGAGCACTGCCTGCACCCCAAGCTGC
AGAGCACCAAGCGCTTCATCAAGTGGTACAACGCCTGGAACGAGAAGCGCAGGGTCTACGAA
GAATAGGGTGAAAAACCTCAGAAGGGAAAACTCCAAACCAGTTGGGAGACTTGTGCAAAGGA
CTTTGCAGATTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCCTTTC
TTTCTCACAGGCATAAGACACAAATTATATATTGTTATGAAGCACTTTTTACCAACGGTCAG
TTTTTACATTTTATAGCTGCGTGCGAAAGGCTTCCAGATGGGAGACCCATCTCTCTTGTGCT
CCAGACTTCATCACAGGCTGCTTTTTATCAAAAAGGGGAAAACTCATGCCTTTCCTTTTTAA
AAAATGCTTTTTTGTATTTGTCCATACGTCACTATACATCTGAGCTTTATAAGCGCCCGGGA
GGAACAATGAGCTTGGTGGACACATTTCATTGCAGTGTTGCTCCATTCCTAGCTTGGGAAGC
TTCCGCTTAGAGGTCCTGGCGCCTCGGCACAGCTGCCACGGGCTCTCCTGGGCTTATGGCCG
GTCACAGCCTCAGTGTGACTCCACAGTGGCCCCTGTAGCCGGGCAAGCAGGAGCAGGTCTCT
CTGCATCTGTTCTCTGAGGAACTCAAGTTTGGTTGCCAGAAAAATGTGCTTCATTCCCCCCT
GGTTAATTTTTACACACCCTAGGAAACATTTCCAAGATCCTGTGATGGCGAGACAAATGATC
CTTAAAGAAGGTGTGGGGTCTTTCCCAACCTGAGGATTTCTGAAAGGTTCACAGGTTCAATA
TTTAATGCTTCAGAAGCATGTGAGGTTCCCAACACTGTCAGCAAAAACCTTAGGAGAAAACT
TAAAAATATATGAATACATGCGCAATACACAGCTACAGACACACATTCTGTTGACAAGGGAA
AACCTTCAAAGCATGTTTCTTTCCCTCACCACAACAGAACATGCAGTACTAAAGCAATATAT
TTGTGATTCCCCATGTAATTCTTCAATGTTAAACAGTGCAGTCCTCTTTCGAAAGCTAAGAT
GACCATGCGCCCTTTCCTCTGTACATATACCCTTAAGAACGCCCCCTCCACACACTGCCCCC
CAGTATATGCCGCATTGTACTGCTGTGTTATATGCTATGTACATGTCAGAAACCATTAGCAT
TGCATGCAGGTTTCATATTCTTTCTAAGATGGAAAGTAATAAAATATATTTGAAATGTAAAA
AAAAAAAAAAA

FIGURE 149

MSLLPRRAPPVSMRLLAAALLLLLLALYTARVDGSKCKCSRKGPKIRYSDVKKLEMKPKYPH
CEEKMVIITTKSVSRYRGQEHCLHPKLQSTKRFIKWYNAWNEKRRVYEE

FIGURE 150

GCCCCAGGGACTGCTATGGCTTCCTTTGTTGTTCACCCCGGTCTGCGTCATGTTAAACTCCA
ATGTCCTCCTGTGGTTAACTGCTCTTGCCATCAAGTTCACCCTCATTGACAGCCAAGCACAG
TATCCAGTTGTCAACACAAATTATGGCAAAATCCGGGGCCTAAGAACACCGTTACCCAATGA
GATCTTGGGTCCAGTGGAGCAGTACTTAGGGGTCCCCTATGCCTCACCCCCCACTGGAGAGA
GGCGGTTTCAGCCCCCAGAACCCCCGTCCTCCTGGACTGGCATCCGAAATACTACTCAGTTT
GCTGCTGTGTGCCCCCAGCACCTGGATGAGAGATCCTTACTGCATGACATGCTGCCCATCTG
GTTTACCGCCAATTTGGATACTTTGATGACCTATGTTCAAGATCAAAATGAAGACTGCCTTT
ACTTAAACATCTACGTGCCCACGGAAGATGGAGCCAACACAAAGAAAAACGCAGATGATATA
ACGAGTAATGACCGTGGTGAAGACGAAGATATTCATGATCAGAACAGTAAGAAGCCCGTCAT
GGTCTATATCCATGGGGGATCTTACATGGAGGGCACCGGCAACATGATTGACGGCAGCATTT
TGGCAAGCTACGGAAACGTCATCGTGATCACCATTAACTACCGTCTGGGAATACTAGGGTTT
TTAAGTACCGGTGACCAGGCAGCAAAAGGCAACTATGGGCTCCTGGATCAGATTCAAGCACT
GCGGTGGATTGAGGAGAATGTGGGAGCCTTTGGCGGGGACCCCAAGAGAGTGACCATCTTTG
GCTCGGGGGCTGGGGCCTCCTGTGTCAGCCTGTTGACCCTGTCCCACTACTCAGAAGGTCTC
TTCCAGAAGGCCATCATTCAGAGCGGCACCGCCCTGTCCAGCTGGGCAGTGAACTACCAGCC
GGCCAAGTACACTCGGATATTGGCAGACAAGGTCGGCTGCAACATGCTGGACACCACGGACA
TGGTAGAATGCCTGCGGAACAAGAACTACAAGGAGCTCATCCAGCAGACCATCACCCCGGCC
ACCTACCACATAGCCTTCGGGCCGGTGATCGACGGCGACGTCATCCCAGACGACCCCCAGAT
CCTGATGGAGCAAGGCGAGTTCCTCAACTACGACATCATGCTGGGCGTCAACCAAGGGGAAG
GCCTGAAGTTCGTGGACGGCATCGTGGATAACGAGGACGGTGTGACGCCCAACGACTTTGAC
TTCTCCGTGTCCAACTTCGTGGACAACCTTTACGGCTACCCTGAAGGGAAAGACACTTTGCG
GGAGACTATCAAGTTCATGTACACAGACTGGGCCGATAAGGAAAACCCGGAGACGCGGCGGA
AAACCCTGGTGGCTCTCTTTACTGACCACCAGTGGGTGGCCCCCGCCGTGGCCGCCGACCTG
CACGCGCAGTACGGCTCCCCCACCTACTTCTATGCCTTCTATCATCACTGCCAAAGCGAAAT
GAAGCCCAGCTGGGCAGATTCGGCCCATGGTGATGAGGTCCCCTATGTCTTCGGCATCCCCA
TGATCGGTCCCACCGAGCTCTTCAGTTGTAACTTTTCCAAGAACGACGTCATGCTCAGCGCC
GTGGTCATGACCTACTGGACGAACTTCGCCAAAACTGGTGATCCAAATCAACCAGTTCCTCA
GGATACCAAGTTCATTCACACAAAACCCAACCGCTTTGAAGAAGTGGCCTGGTCCAAGTATA
ATCCCAAAGACCAGCTCTATCTGCATATTGGCTTGAAACCCAGAGTGAGAGATCACTACCGG
GCAACGAAAGTGGCTTTCTGGTTGGAACTCGTTCCTCATTTGCACAACTTGAACGAGATATT
CCAGTATGTTTCAACAACCACAAAGGTTCCTCCACCAGACATGACATCATTTCCCTATGGCA
CCCGGCGATCTCCCGCCAAGATATGGCCAACCACCAAACGCCCAGCAATCACTCCTGCCAAC
AATCCCAAACACTCTAAGGACCCTCACAAAACAGGGCCTGAGGACACAACTGTCCTCATTGA
AACCAAACGAGATTATTCCACCGAATTAAGTGTCACCATTGCCGTCGGGGCGTCGCTCCTCT
TCCTCAACATCTTAGCTTTTGCGGCGCTGTACTACAAAAAGGACAAGAGGCGCCATGAGACT
CACAGGCGCCCCAGTCCCCAGAGAAACACCACAAATGATATCGCTCACATCCAGAACGAAGA
GATCATGTCTCTGCAGATGAAGCAGCTGGAACACGATCACGAGTGTGAGTCGCTGCAGGCAC
ACGACACACTGAGGCTCACCTGCCCGCCAGACTACACCCTCACGCTGCGCCGGTCGCCAGAT
GACATCCCACTTATGACGCCAAACACCATCACCATGATTCCAAACACACTGACGGGGATGCA
GCCTTTGCACACTTTTAACACCTTCAGTGGAGGACAAAACAGTACAAATTTACCCCACGGAC
ATTCCACCACTAGAGTATAGCTTTGCCCTATTTCCCTTCCTATCCCTCTGCCCTACCCGCTC
AGCAACATAGAAGAGGGAAGGAAAGAGAGAAGGAAAGAGAGAGAGAAAGAAAGTCTCCAGAC
CAGGAATGTTTTTGTCCCACTGACTTAAGACAAAAATGCAAAAAGGCAGTCATCCCATCCCG
GCAGACCCTTATCGTTGGTGTTTTCCAGTATTACAAGATCAACTTCTGACCCTGTGAAATGT
GAGAAGTACACATTTCTGTTAAAATAACTGCTTTAAGATCTCTACCACTCCAATCAATGTTT
AGTGTGATAGGACATCACCATTTCAAGGCCCCGGGTGTTTCCAACGTCATGGAAGCAGCTGA
CACTTCTGAAACTCAGCCAAGGACACTTGATATTTTTTAATTACAATGGAAGTTTAAACATT
TCTTTCTGTGCCACACAATGGATGGCTCTCCTTAAGTGAAGAAAGAGTCAATGAGATTTTGC
CCAGCACATGGAGCTGTAATCCAGAGAGAAGGAAACGTAGAAATTTATTATTAAAAGAATGG
ACTGTGCAGCGAAATCTGTACGGTTCTGTGCAAAGAGGTGTTTTGCCAGCCTGAACTATATT
TAAGAGACTTTGT

FIGURE 151

MLNSNVLLWLTALAIKFTLIDSQAQYPVVNTNYGKIRGLRTPLPNEILGPVEQYLGVPYASP
PTGERRFQPPEPPSSWTGIRNTTQFAAVCPQHLDERSLLHDMLPIWFTANLDTLMTYVQDQN
EDCLYLNIYVPTEDGANTKKNADDITSNDRGEDEDIHDQNSKKPVMVYIHGGSYMEGTGNMI
DGSILASYGNVIVITINYRLGILGFLSTGDQAAKGNYGLLDQIQALRWIEENVGAFGGDPKR
VTIFGSGAGASCVSLLTLSHYSEGLFQKAIIQSGTALSSWAVNYQPAKYTRILADKVGCNML
DTTDMVECLRNKNYKELIQQTITPATYHIAFGPVIDGDVIPDDPQILMEQGEFLNYDIMLGV
NQGEGLKFVDGIVDNEDGVTPNDFDFSVSNFVDNLYGYPEGKDTLRETIKFMYTDWADKENP
ETRRKTLVALFTDHQWVAPAVAADLHAQYGSPTYFYAFYHHCQSEMKPSWADSAHGDEVPYV
FGIPMIGPTELFSCNFSKNDVMLSAVVMTYWTNFAKTGDPNQPVPQDTKFIHTKPNRFEEVA
WSKYNPKDQLYLHIGLKPRVRDHYRATKVAFWLELVPHLHNLNEIFQYVSTTTKVPPPDMTS
FPYGTRRSPAKIWPTTKRPAITPANNPKHSKDPHKTGPEDTTVLIETKRDYSTELSVTIAVG
ASLLFLNILAFAALYYKKDKRRHETHRRPSPQRNTTNDIAHIQNEEIMSLQMKQLEHDHECE
SLQAHDTLRLTCPPDYTLTLRRSPDDIPLMTPNTITMIPNTLTGMQPLHTFNTFSGGQNSTN
LPHGHSTTRV

FIGURE 152

GGGAAAGATGGCGGCGACTCTGGGACCCCTTGGGTCGTGGCAGCAGTGGCGGCGATGTTTGT
CGGCTCGGGATGGGTCCAGGATGTTACTCCTTCTTCTTTTGTTGGGGTCTGGGCAGGGGCCA
CAGCAAGTCGGGGCGGGTCAAACGTTCGAGTACTTGAAACGGGAGCACTCGCTGTCGAAGCC
CTACCAGGGTGTGGGCACAGGCAGTTCCTCACTGTGGAATCTGATGGGCAATGCCATGGTGA
TGACCCAGTATATCCGCCTTACCCCAGATATGCAAAGTAAACAGGGTGCCTTGTGGAACCGG
GTGCCATGTTTCCTGAGAGACTGGGAGTTGCAGGTGCACTTCAAAATCCATGGACAAGGAAA
GAAGAATCTGCATGGGGATGGCTTGGCAATCTGGTACACAAAGGATCGGATGCAGCCAGGGC
CTGTGTTTGGAAACATGGACAAATTTGTGGGGCTGGGAGTATTTGTAGACACCTACCCCAAT
GAGGAGAAGCAGCAAGAGCGGGTATTCCCCTACATCTCAGCCATGGTGAACAACGGCTCCCT
CAGCTATGATCATGAGCGGGATGGGCGGCCTACAGAGCTGGGAGGCTGCACAGCCATTGTCC
GCAATCTTCATTACGACACCTTCCTGGTGATTCGCTACGTCAAGAGGCATTTGACGATAATG
ATGGATATTGATGGCAAGCATGAGTGGAGGGACTGCATTGAAGTGCCCGGAGTCCGCCTGCC
CCGCGGCTACTACTTCGGCACCTCCTCCATCACTGGGGATCTCTCAGATAATCATGATGTCA
TTTCCTTGAAGTTGTTTGAACTGACAGTGGAGAGAACCCCAGAAGAGGAAAAGCTCCATCGA
GATGTGTTCTTGCCCTCAGTGGACAATATGAAGCTGCCTGAGATGACAGCTCCACTGCCGCC
CCTGAGTGGCCTGGCCCTCTTCCTCATCGTCTTTTTCTCCCTGGTGTTTTCTGTATTTGCCA
TAGTCATTGGTATCATACTCTACAACAAATGGCAGGAACAGAGCCGAAAGCGCTTCTACTGA
GCCCTCCTGCTGCCACCACTTTTGTGACTGTCACCCATGAGGTATGGAAGGAGCAGGCACTG
GCCTGAGCATGCAGCCTGGAGAGTGTTCTTGTCTCTAGCAGCTGGTTGGGGACTATATTCTG
TCACTGGAGTTTTGAATGCAGGGACCCCGCATTCCCATGGTTGTGCATGGGGACATCTAACT
CTGGTCTGGGAAGCCACCCACCCCAGGGCAATGCTGCTGTGATGTGCCTTTCCCTGCAGTCC
TTCCATGTGGGAGCAGAGGTGTGAAGAGAATTTACGTGGTTGTGATGCCAAAATCACAGAAC
AGAATTTCATAGCCCAGGCTGCCGTGTTGTTTGACTCAGAAGGCCCTTCTACTTCAGTTTTG
AATCCACAAAGAATTAAAAACTGGTAACACCACAGGCTTTCTGACCATCCATTCGTTGGGTT
TTGCATTTGACCCAACCCTCTGCCTACCTGAGGAGCTTTCTTTGGAAACCAGGATGGAAACT
TCTTCCCTGCCTTACCTTCCTTTCACTCCATTCATTGTCCTCTCTGTGTGCAACCTGAGCTG
GGAAAGGCATTTGGATGCCTCTCTGTTGGGGCCTGGGGCTGCAGAACACACCTGCGTTTCAC
TGGCCTTCATTAGGTGGCCCTAGGGAGATGGCTTTCTGCTTTGGATCACTGTTCCCTAGCAT
GGGTCTTGGGTCTATTGGCATGTCCATGGCCTTCCCAATCAAGTCTCTTCAGGCCCTCAGTG
AAGTTTGGCTAAAGGTTGGTGTAAAAATCAAGAGAAGCCTGGAAGACATCATGGATGCCATG
GATTAGCTGTGCAACTGACCAGCTCCAGGTTTGATCAAACCAAAAGCAACATTTGTCATGTG
GTCTGACCATGTGGAGATGTTTCTGGACTTGCTAGAGCCTGCTTAGCTGCATGTTTTGTAGT
TACGATTTTTGGAATCCCACTTTGAGTGCTGAAAGTGTAAGGAAGCTTTCTTCTTACACCTT
GGGCTTGGATATTGCCCAGAGAAGAAATTTGGCTTTTTTTTTCTTAATGGACAAGAGACAGT
TGCTGTTCTCATGTTCCAAGTCTGAGAGCAACAGACCCTCATCATCTGTGCCTGGAAGAGTT
CACTGTCATTGAGCAGCACAGCCTGAGTGCTGGCCTCTGTCAACCCTTATTCCACTGCCTTA
TTTGACAAGGGGTTACATGCTGCTCACCTTACTGCCCTGGGATTAAATCAGTTACAGGCCAG
AGTCTCCTTGGAGGGCCTGGAACTCTGAGTCCTCCTATGAACCTCTGTAGCCTAAATGAAAT
TCTTAAAATCACCGATGGAACCAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCG
ACCTGCAGTAGGGATAACAGGGTAATAAGCTTGGCCGCCATGG

FIGURE 153

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50911

><subunit 1 of 1, 348 aa, 1 stop

><MW: 39711, pI: 8.70, NX(S/T): 1

MAATLGPLGSWQQWRRCLSARDGSRMLLLLLLLGSGQGPQQVGAGQTFEYLKREHSLSKPYQ
GVGTGSSSLWNLMGNAMVMTQYIRLTPDMQSKQGALWNRVPCFLRDWELQVHFKIHGQGKKN
LHGDGLAIWYTKDRMQPGPVFGNMDKFVGLGVFVDTYPNEEKQQERVFPYISAMVNNGSLSY
DHERDGRPTELGGCTAIVRNLHYDTFLVIRYVKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG
YYFGTSSITGDLSDNHDVISLKLFELTVERTPEEEKLHRDVFLPSVDNMKLPEMTAPLPPLS
GLALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRKRFY

FIGURE 154

CCGAGCCGGGCGCGCAGCGACGGAGCTGGGGCCGGCCTGGGACCATGGGCGTGAGTGCAATC
TACGGATCAGTCTCTGATGGTGGGTCGTTAACCTCAGTGGGGACTCCAAGATTTCCATGAAG
AAAATCAGTTGTCTTCATTCAAGAATTGGGGTCTGGCTCAGAATTCCTGCAGCTGGTGAAAA
TCTGTTTTCTAGAAGAGGTTTAATTAATGCCTGCAGTCTGACATGTTCCCGATTTGAGGTGA
AACCATGAAGAGAAAATAGAATACTTAATA**ATG**CTTTTCCGCAACCGCTTCTTGCTGCTGCT
GGCCCTGGCTGCGCTGCTGGCCTTTGTGAGCCTCAGCCTGCAGTTCTTCCACCTGATCCCGG
TGTCGACTCCTAAGAATGGAATGAGTAGCAAGAGTCGAAAGAGAATCATGCCCGACCCTGTG
ACGGAGCCCCCTGTGACAGACCCCGTTTATGAAGCTCTTTTGTACTGCAACATCCCCAGTGT
GGCCGAGCGCAGCATGGAAGGTCATGCCCCGCATCATTTTAAGCTGGTCTCAGTGCATGTGT
TCATTCGCCACGGAGACAGGTACCCACTGTATGTCATTCCCAAAACAAAGCGACCAGAAATT
GACTGCACTCTGGTGGCTAACAGGAAACCGTATCACCCAAAACTGGAAGCTTTCATTAGTCA
CATGTCAAAAGGATCCGGAGCCTCTTTCGAAAGCCCCTTGAACTCCTTGCCTCTTTACCCAA
ATCACCCATTGTGTGAGATGGGAGAGCTCACACAGACAGGAGTTGTGCAGCATTTGCAGAAC
GGTCAGCTGCTGAGGGATATCTATCTAAAGAAACACAAACTCCTGCCCAATGATTGGTCTGC
AGACCAGCTCTATTTAGAGACCACTGGGAAAAGCCGGACCCTACAAAGTGGGCTGGCCTTGC
TTTATGGCTTTCTCCCAGATTTTGACTGGAAGAAGATTTATTTCAGGCACCAGCCAAGTGCG
CTGTTCTGCTCTGGAAGCTGCTATTGCCCGGTAAGAAACCAGTATCTGGAAAAGGAGCAGCG
TCGTCAGTACCTCCTACGTTTGAAAAACAGCCAGCTGGAGAAGACCTACGGGGAGATGGCCA
AGATCGTGGATGTCCCCACCAAGCAGCTTAGAGCTGCCAACCCCATAGACTCCATGCTCTGC
CACTTCTGCCACAATGTCAGCTTTCCCTGTACCAGAAATGGCTGTGTTGACATGGAGCACTT
CAAGGTAATTAAGACCCATCAGATCGAGGATGAAAGGGAAAGACGGGAGAAGAAATTGTACT
TCGGGTATTCTCTCCTGGGTGCCCACCCCATCCTGAACCAAACCATCGGCCGGATGCAGCGT
GCCACCGAGGGCAGGAAAGAAGAGCTCTTTGCCCTCTACTCTGCTCATGATGTCACTCTGTC
ACCAGTTCTCAGTGCCTTGGGCCTTTCAGAAGCCAGGTTCCCAAGGTTTGCAGCCAGGTTGA
TCTTTGAGCTTTGGCAAGACAGAGAAAAGCCCAGTGAACATTCCGTCCGGATTCTTTACAAT
GGCGTCGATGTCACATTCCACACCTCTTTCTGCCAAGACCACCACAAGCGTTCTCCCAAGCC
CATGTGCCCGCTTGAAAACTTGGTCCGCTTTGTGAAAAGGGACATGTTTGTAGCCCTGGGTG
GCAGTGGTACAAATTATTATGATGCATGTCACAGGGAAGGATTC**TAA**AAGGTATGCAGTACA
GCAGTATAGAATCCATGCCAATACAGAGCATAGGGAAAGGTCCACTTCTAGTTTTGTCTGTT
ACTAAGGGTAGAAGATTATTGCTTTTTAAAGGCTAAATATTGTTTGTGGGAACCACAGATGG
TTGGGGTTGAACAGTAAGCACATTGCTGCAATGTGGTACGTGAATTGCTTGGTACAAAATGG
CCAGTTCACAGAGGAATAGAAGGTACTTTATCATAGCCAGACTTCGCTTAGAATGCCAGAAT
AATATAGTTCAAGACCTGAAGTTGCCAATCCAAGTTTGCACTCTTCTGGCCTGCCCCATGTT
ACTATGTGATGGAACCAGCACACCTCAACCAAAATTTTTTTAATCTTAGACATTTTTACCTT
GTCCTTGTTAAGAATTTCTTGAAGTGATTTATCTAAAATAAAGGTTGGCAAACTTTTTCTGT
AAAGGGCCAGATTGTAAATATTTCAGACTGTGTGGACCAAAAGGCCACATACAGTCTCTGTC
ATAACTACTCAACTCTGTTTCTGAAGCAGGAAAGCCACCACAGACAGTACATAAAGGAATAT
GTGTAGCTGGGTTCCCAGGCCAGACAAAACAGATGGTGACCAGACTTGGCCCCTGGGCTGTA
GTTTGCTGACCCCTCATCTAAAAAATAGGCTATACTACAATTGCACTTCCAGCACTTTGAGA
ACGAGTTGAATACCAAGAATTATTCAATGGTTCCTCCAGTAACTTCTGCTAGAAACACAGAA
TTTGGTCTGTATCTGACACTAGAACAAAACTTGAGGGTAAATAAACATTGAATTAGAATGAA
TCATAGAAAACTGATTAGAAGAATACTTGATGTTTATGATGATTGTGGTACAAGATAGTTTT
AAGTATGTTCTAAATATTTGTCTGCTGTAGTCTATTTGCTGTATATGCTGAAATTTTTGTAT
GCCATTTAGTATTTTTATAGTTTAGGAAAATATTTTCTAAGACCAGTTTTAGATGACTCTTA
TTCCTGTAGTAATATTCAATTTGCTGTACCTGCTTGGTGGTTAGAAGGAGGCTAGAAGATGA
ATTCAGGCACTTTCTTCCAATAAAACTAATTATGGCTCATTCCCTTTGACAAGCTGTAGAAC
TGGATTCATTTTTAAACCATTTTCATCAGTTTCAAATGGTAAATTCTGATTGATTTTTAAAT
GCGTTTTTGGAAGAACTTTGCTATTAGGTAGTTTACAGATCTTTATAAGGTGTTTTATATAT
TAGAAGCAATTATAATTACATCTGTGATTTCTGAACTAATGGTGCTAATTCAGAGAAATGGA
AAGTGAAAGTGAGATTCTCTGTTGTCATCGGCATTCCAACTTTTTCTCTTTGTTTTTGTCCA
GTGTTGCATTTGAATATGTCTGTTTCTATAAATAAATTTTTTAAGAATAA

FIGURE 155

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48329

><subunit 1 of 1, 480 aa, 1 stop

><MW: 55240, pI: 9.30, NX(S/T): 2

MLFRNRFLLLLALAALLAFVSLSLQFFHLIPVSTPKNGMSSKSRKRIMPDPVTEPPVTDPVY
EALLYCNIPSVAERSMEGHAPHHFKLVSVHVFIRHGDRYPLYVIPKTKRPEIDCTLVANRKP
YHPKLEAFISHMSKGSGASFESPLNSLPLYPNHPLCEMGELTQTGVVQHLQNGQLLRDIYLK
KHKLLPNDWSADQLYLETTGKSRTLQSGLALLYGFLPDFDWKKIYFRHQPSALFCSGSCYCP
VRNQYLEKEQRRQYLLRLKNSQLEKTYGEMAKIVDVPTKQLRAANPIDSMLCHFCHNVSFPC
TRNGCVDMEHFKVIKTHQIEDERERREKKLYFGYSLLGAHPILNQTIGRMQRATEGRKEELF
ALYSAHDVTLSPVLSALGLSEARFPRFAARLIFELWQDREKPSEHSVRILYNGVDVTFHTSF
CQDHHKRSPKPMCPLENLVRFVKRDMFVALGGSGTNYYDACHREGF

FIGURE 156

AAAAAAGCTCACTAAAGTTTCTATTAGAGCGAATACGGTAGATTTCCATCCCCTTTTGAAGA
ACAGTACTGTGGAGCTATTTAAGAGATAAAAACGAAATATCCTTTCTGGGAGTTCAAGATTG
TGCAGTAATTGGTTAGGACTCTGAGCGCCGCTGTTCACCAATCGGGGAGAGAAAAGCGGAGA
TCCTGCTCGCCTTGCACGCGCCTGAAGCACAAAGCAGATAGCTAGGAATGAACCATCCCTGG
GAGTATGTGGAAACAACGGAGGAGCTCTGACTTCCCAACTGTCCCATTCTATGGGCGAAGGA
ACTGCTCCTGACTTCAGTGGTTAAGGGCAGAATTGAAAATAATTCTGGAGGAAGATAAGAAT
GATTCCTGCGCGACTGCACCGGGACTACAAAGGGCTTGTCCTGCTGGGAATCCTCCTGGGGA
CTCTGTGGGAGACCGGATGCACCCAGATACGCTATTCAGTTCCGGAAGAGCTGGAGAAAGGC
TCTAGGGTGGGCGACATCTCCAGGGACCTGGGGCTGGAGCCCCGGGAGCTCGCGGAGCGCGG
AGTCCGCATCATCCCCAGAGGTAGGACGCAGCTTTTCGCCCTGAATCCGCGCAGCGGCAGCT
TGGTCACGGCGGGCAGGATAGACCGGGAGGAGCTCTGTATGGGGGCCATCAAGTGTCAATTA
AATCTAGACATTCTGATGGAGGATAAAGTGAAAATATATGGAGTAGAAGTAGAAGTAAGGGA
CATTAACGACAATGCGCCTTACTTTCGTGAAAGTGAATTAGAAATAAAAATTAGTGAAAATG
CAGCCACTGAGATGCGGTTCCCTCTACCCCACGCCTGGGATCCGGATATCGGGAAGAACTCT
CTGCAGAGCTACGAGCTCAGCCCGAACACTCACTTCTCCCTCATCGTGCAAAATGGAGCCGA
CGGTAGTAAGTACCCCGAATTGGTGCTGAAACGCGCCCTGGACCGCGAAGAAAAGGCTGCTC
ACCACCTGGTCCTTACGGCCTCCGACGGGGGCGACCCGGTGCGCACAGGCACCGCGCGCATC
CGCGTGATGGTTCTGGATGCGAACGACAACGCACCAGCGTTTGCTCAGCCCGAGTACCGCGC
GAGCGTTCCGGAGAATCTGGCCTTGGGCACGCAGCTGCTTGTAGTCAACGCTACCGACCCTG
ACGAAGGAGTCAATGCGGAAGTGAGGTATTCCTTCCGGTATGTGGACGACAAGGCGGCCCAA
GTTTTCAAACTAGATTGTAATTCAGGGACAATATCAACAATAGGGGAGTTGGACCACGAGGA
GTCAGGATTCTACCAGATGGAAGTGCAAGCAATGGATAATGCAGGATATTCTGCGCGAGCCA
AAGTCCTGATCACTGTTCTGGACGTGAACGACAATGCCCCAGAAGTGGTCCTCACCTCTCTC
GCCAGCTCGGTTCCCGAAAACTCTCCCAGAGGGACATTAATTGCCCTTTTAAATGTAAATGA
CCAAGATTCTGAGGAAAACGGACAGGTGATCTGTTTCATCCAAGGAAATCTGCCCTTTAAAT
TAGAAAAATCTTACGGAAATTACTATAGTTTAGTCACAGACATAGTCTTGGATAGGGAACAG
GTTCCTAGCTACAACATCACAGTGACCGCCACTGACCGGGGAACCCCGCCCCTATCCACGGA
AACTCATATCTCGCTGAACGTGGCAGACACCAACGACAACCCGCCGGTCTTCCCTCAGGCCT
CCTATTCCGCTTATATCCCAGAGAACAATCCCAGAGGAGTTTCCCTCGTCTCTGTGACCGCC
CACGACCCCGACTGTGAAGAGAACGCCCAGATCACTTATTCCCTGGCTGAGAACACCATCCA
AGGGGCAAGCCTATCGTCCTACGTGTCCATCAACTCCGACACTGGGGTACTGTATGCGCTGA
GCTCCTTCGACTACGAGCAGTTCCGAGACTTGCAAGTGAAAGTGATGGCGCGGGACAACGGG
CACCCGCCCCTCAGCAGCAACGTGTCGTTGAGCCTGTTCGTGCTGGACCAGAACGACAATGC
GCCCGAGATCCTGTACCCCGCCCTCCCCACGGACGGTTCCACTGGCGTGGAGCTGGCTCCCC
GCTCCGCAGAGCCCGGCTACCTGGTGACCAAGGTGGTGGCGGTGGACAGAGACTCCGGCCAG
AACGCCTGGCTGTCCTACCGTCTGCTCAAGGCCAGCGAGCCGGGACTCTTCTCGGTGGGTCT
GCACACGGGCGAGGTGCGCACGGCGCGAGCCCTGCTGGACAGAGACGCGCTCAAGCAGAGCC
TCGTAGTGGCCGTCCAGGACCACGGCCAGCCCCCTCTCTCCGCCACTGTCACGCTCACCGTG
GCCGTGGCCGACAGCATCCCCCAAGTCCTGGCGGACCTCGGCAGCCTCGAGTCTCCAGCTAA
CTCTGAAACCTCAGACCTCACTCTGTACCTGGTGGTAGCGGTGGCCGCGGTCTCCTGCGTCT
TCCTGGCCTTCGTCATCTTGCTGCTGGCGCTCAGGCTGCGGCGCTGGCACAAGTCACGCCTG
CTGCAGGCTTCAGGAGGCGGCTTGACAGGAGCGCCGGCGTCGCACTTTGTGGGCGTGGACGG
GGTGCAGGCTTTCCTGCAGACCTATTCCCACGAGGTTTCCCTCACCACGGACTCGCGGAAGA
GTCACCTGATCTTCCCCCAGCCCAACTATGCAGACATGCTCGTCAGCCAGGAGAGCTTTGAA
AAAAGCGAGCCCCTTTTGCTGTCAGGTGATTCGGTATTTTCTAAAGACAGTCATGGGTTAAT
TGAGGTGAGTTTATATCAAATCTTCTTTCTTTTTTTTTTTAATTGCTCTGTCTCCCAAGCTG
GAGTGCAGCGGTACGATCATAGCTCACTGCGGCCTCAAACTCCTAGGCTCAAGCAATTATCC
CACCTTTGCCTCCGGTGTAACAGGGACTACAGGTGCAAGCCACCTACTGTCTGCCTATCTAT
CTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATTACTTTCTTGTACAGACG
GGAGTCTCACGCCTGTAATCCCAGTACTTTGGGAGGCCGAGGCGGGTGGATCACCTGAGGTT
GGGAGTTTGAGACCAGCCTGACCAACATGGAGAAACCCCGTCTATACTAAAAAAATACAAAA
TTAGCCGGGCGTGGTGGTGCATGTCTGTAATCCCAGCTACTTGGGAGGCTGAGTCAGGAGAA
TTGCTTTAACCTGGGAGGTGGAGGTTGCAATGAGCTGAGATTGTGCCATTGCACTCCAGCCT
GGGCAACAAGAGTGAAACTCTATCTCA

FIGURE 157

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48306
><subunit 1 of 1, 916 aa, 1 stop
><MW: 100204, pI: 4.92, NX(S/T): 4

MIPARLHRDYKGLVLLGILLGTLWETGCTQIRYSVPEELEKGSRVGDISRDLGLEPRELAER
GVRIIPRGRTQLFALNPRSGSLVTAGRIDREELCMGAIKCQLNLDILMEDKVKIYGVEVEVR
DINDNAPYFRESELEIKISENAATEMRFPLPHAWDPDIGKNSLQSYELSPNTHFSLIVQNGA
DGSKYPELVLKRALDREEKAAHHLVLTASDGGDPVRTGTARIRVMVLDANDNAPAFAQPEYR
ASVPENLALGTQLLVVNATDPDEGVNAEVRYSFRYVDDKAAQVFKLDCNSGTISTIGELDHE
ESGFYQMEVQAMDNAGYSARAKVLITVLDVNDNAPEVVLTSLASSVPENSPRGTLIALLNVN
DQDSEENGQVICFIQGNLPFKLEKSYGNYYSLVTDIVLDREQVPSYNITVTATDRGTPPLST
ETHISLNVADTNDNPPVFPQASYSAYIPENNPRGVSLVSVTAHDPDCEENAQITYSLAENTI
QGASLSSYVSINSDTGVLYALSSFDYEQFRDLQVKVMARDNGHPPLSSNVSLSLFVLDQNDN
APEILYPALPTDGSTGVELAPRSAEPGYLVTKVVAVDRDSGQNAWLSYRLLKASEPGLFSVG
LHTGEVRTARALLDRDALKQSLVVAVQDHGQPPLSATVTLTVAVADSIPQVLADLGSLESPA
NSETSDLTLYLVVAVAAVSCVFLAFVILLLALRLRRWHKSRLLQASGGGLTGAPASHFVGVD
GVQAFLQTYSHEVSLTTDSRKSHLIFPQPNYADMLVSQESFEKSEPLLLSGDSVFSKDSHGL
IEVSLYQIFFLFFFNCSVSQAGVQRYDHSSLRPQTPRLKQLSHLCLRCNRDYRCKPPTVCLS
IYLSIYLSIYLSIYLLLSCTDGSLTPVIPVLWEAEAGGSPEVGSLRPA

FIGURE 158

CCCAGGCTCTAGTGCAGGAGGAGAAGGAGGAGGAGCAGGAGGTGGAGATTCCCAGTTAAAAG
GCTCCAGAATCGTGTACCAGGCAGAGAACTGAAGTACTGGGGCCTCCTCCACTGGGTCCGAA
TCAGTAGGTGACCCCGCCCCTGGATTCTGGAAGACCTCACC**ATG**GGACGCCCCCGACCTCGT
GCGGCCAAGACGTGGATGTTCCTGCTCTTGCTGGGGGGAGCCTGGGCAGGACACTCCAGGGC
ACAGGAGGACAAGGTGCTGGGGGGTCATGAGTGCCAACCCCATTCGCAGCCTTGGCAGGCGG
CCTTGTTCCAGGGCCAGCAACTACTCTGTGGCGGTGTCCTTGTAGGTGGCAACTGGGTCCTT
ACAGCTGCCCACTGTAAAAAACCGAAATACACAGTACGCCTGGGAGACCACAGCCTACAGAA
TAAAGATGGCCCAGAGCAAGAAATACCTGTGGTTCAGTCCATCCCACACCCCTGCTACAACA
GCAGCGATGTGGAGGACCACAACCATGATCTGATGCTTCTTCAACTGCGTGACCAGGCATCC
CTGGGGTCCAAAGTGAAGCCCATCAGCCTGGCAGATCATTGCACCCAGCCTGGCCAGAAGTG
CACCGTCTCAGGCTGGGGCACTGTCACCAGTCCCCGAGAGAATTTTCCTGACACTCTCAACT
GTGCAGAAGTAAAAATCTTTCCCCAGAAGAAGTGTGAGGATGCTTACCCGGGGCAGATCACA
GATGGCATGGTCTGTGCAGGCAGCAGCAAAGGGGCTGACACGTGCCAGGGCGATTCTGGAGG
CCCCCTGGTGTGTGATGGTGCACTCCAGGGCATCACATCCTGGGGCTCAGACCCCTGTGGGA
GGTCCGACAAACCTGGCGTCTATACCAACATCTGCCGCTACCTGGACTGGATCAAGAAGATC
ATAGGCAGCAAGGGC**TGA**TTCTAGGATAAGCACTAGATCTCCCTTAATAAACTCACAACTCT
CTGGTTC

FIGURE 159

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48336
<subunit 1 of 1, 260 aa, 1 stop
<MW: 28048, pI: 7.87, NX(S/T): 1
MGRPRPRAAKTWMFLLLLGGAWAGHSRAQEDKVLGGHECQPHSQPWQAALFQGQQLLCGGVL
VGGNWVLTAAHCKKPKYTVRLGDHSLQNKDGPEQEIPVVQSIPHPCYNSSDVEDHNHDLMLL
QLRDQASLGSKVKPISLADHCTQPGQKCTVSGWGTVTSPRENFPDTLNCAEVKIFPQKKCED
AYPGQITDGMVCAGSSKGADTCQGDSGGPLVCDGALQGITSWGSDPCGRSDKPGVYTNICRY
LDWIKKIIGSKG

Important Features:

Signal peptide:
amino acids 1-23

Transmembrane domain:
amino acids 51-71

N-glycosylation site.
amino acids 110-113

Serine proteases, trypsin family, histidine active site.
amino acids 69-74 and 207-217

Tyrosine kinase phosphorylation site.
amino acids 182-188

Kringle domain proteins motif
amino acids 205-217

FIGURE 160

GGCGCCGGTGCACCGGGCGGGCTGAGCGCCTCCTGCGGCCCGGCCTGCGCGCCCCGGCCCGC
CGCGCCGCCCACGCCCCAACCCCGGCCCGCGCCCCCTAGCCCCCGCCCGGGCCCGCGCCCGC
GCCCGCGCCCAGGTGAGCGCTCCGCCCGCCGCGAGGCCCCGCCCCGGCCCGCCCCCGCCCCG
CCCCGGCCGGCGGGGGAACCGGGCGGATTCCTCGCGCGTCAAACCACCTGATCCCATAAAAC
ATTCATCCTCCCGGCGGCCCGCGCTGCGAGCGCCCCGCCAGTCCGCGCCGCCGCCGCCCTCG
CCCTGTGCGCCCTGCGCGCCCTGCGCACCCGCGGCCCGAGCCCAGCCAGAGCCGGGCGGAGC
GGAGCGCGCCGAGCCTCGTCCCGCGGCCGGGCCGGGGCCGGGCCGTAGCGGCGGCGCCTGGA
TGCGGACCCGGCCGCGGGGAGACGGGCGCCCGCCCCGAAACGACTTTCAGTCCCCGACGCGC
CCCGCCCAACCCCTACG**ATG**AAGAGGGCGTCCGCTGGAGGGAGCCGGCTGCTGGCATGGGTG
CTGTGGCTGCAGGCCTGGCAGGTGGCAGCCCCATGCCCAGGTGCCTGCGTATGCTACAATGA
GCCCAAGGTGACGACAAGCTGCCCCCAGCAGGGCCTGCAGGCTGTGCCCGTGGGCATCCCTG
CTGCCAGCCAGCGCATCTTCCTGCACGGCAACCGCATCTCGCATGTGCCAGCTGCCAGCTTC
CGTGCCTGCCGCAACCTCACCATCCTGTGGCTGCACTCGAATGTGCTGGCCCGAATTGATGC
GGCTGCCTTCACTGGCCTGGCCCTCCTGGAGCAGCTGGACCTCAGCGATAATGCACAGCTCC
GGTCTGTGGACCCTGCCACATTCCACGGCCTGGGCCGCCTACACACGCTGCACCTGGACCGC
TGCGGCCTGCAGGAGCTGGGCCCGGGGCTGTTCCGCGGCCTGGCTGCCCTGCAGTACCTCTA
CCTGCAGGACAACGCGCTGCAGGCACTGCCTGATGACACCTTCCGCGACCTGGGCAACCTCA
CACACCTCTTCCTGCACGGCAACCGCATCTCCAGCGTGCCCGAGCGCGCCTTCCGTGGGCTG
CACAGCCTCGACCGTCTCCTACTGCACCAGAACCGCGTGGCCCATGTGCACCCGCATGCCTT
CCGTGACCTTGGCCGCCTCATGACACTCTATCTGTTTGCCAACAATCTATCAGCGCTGCCCA
CTGAGGCCCTGGCCCCCCTGCGTGCCCTGCAGTACCTGAGGCTCAACGACAACCCCTGGGTG
TGTGACTGCCGGGCACGCCCACTCTGGGCCTGGCTGCAGAAGTTCCGCGGCTCCTCCTCCGA
GGTGCCCTGCAGCCTCCCGCAACGCCTGGCTGGCCGTGACCTCAAACGCCTAGCTGCCAATG
ACCTGCAGGGCTGCGCTGTGGCCACCGGCCCTTACCATCCCATCTGGACCGGCAGGGCCACC
GATGAGGAGCCGCTGGGGCTTCCCAAGTGCTGCCAGCCAGATGCCGCTGACAAGGCCTCAGT
ACTGGAGCCTGGAAGACCAGCTTCGGCAGGCAATGCGCTGAAGGGACGCGTGCCGCCCGGTG
ACAGCCCGCCGGGCAACGGCTCTGGCCCACGGCACATCAATGACTCACCCTTTGGGACTCTG
CCTGGCTCTGCTGAGCCCCCGCTCACTGCAGTGCGGCCCGAGGGCTCCGAGCCACCAGGGTT
CCCCACCTCGGGCCCTCGCCGGAGGCCAGGCTGTTCACGCAAGAACCGCACCCGCAGCCACT
GCCGTCTGGGCCAGGCAGGCAGCGGGGGTGGCGGGACTGGTGACTCAGAAGGCTCAGGTGCC
CTACCCAGCCTCACCTGCAGCCTCACCCCCCTGGGCCTGGCGCTGGTGCTGTGGACAGTGCT
TGGGCCCTGC**TGA**CCCCCAGCGGACACAAGAGCGTGCTCAGCAGCCAGGTGTGTGTACATAC
GGGGTCTCTCTCCACGCCGCCAAGCCAGCCGGGCGGCCGACCCGTGGGGCAGGCCAGGCCAG
GTCCTCCCTGATGGACGCCTGCCGCCCGCCACCCCCATCTCCACCCCATCATGTTTACAGGG
TTCGGCGGCAGCGTTTGTTCCAGAACGCCGCCTCCCACCCAGATCGCGGTATATAGAGATAT
GCATTTTATTTTACTTGTGTAAAAATATCGGACGACGTGGAATAAAGAGCTCTTTTCTTAAA
AAAA

FIGURE 161

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44184
><subunit 1 of 1, 473 aa, 1 stop
><MW: 50708, pI: 9.28, NX(S/T): 6
MKRASAGGSRLLAWVLWLQAWQVAAPCPGACVCYNEPKVTTSCPQQGLQAVPVGIPAASQRI
FLHGNRISHVPAASFRACRNLTILWLHSNVLARIDAAAFTGLALLEQLDLSDNAQLRSVDPA
TFHGLGRLHTLHLDRCGLQELGPGLFRGLAALQYLYLQDNALQALPDDTFRDLGNLTHLFLH
GNRISSVPERAFRGLHSLDRLLLHQNRVAHVHPHAFRDLGRLMTLYLFANNLSALPTEALAP
LRALQYLRLNDNPWVCDCRARPLWAWLQKFRGSSSEVPCSLPQRLAGRDLKRLAANDLQGCA
VATGPYHPIWTGRATDEEPLGLPKCCQPDAADKASVLEPGRPASAGNALKGRVPPGDSPPGN
GSGPRHINDSPFGTLPGSAEPPLTAVRPEGSEPPGFPTSGPRRRPGCSRKNRTRSHCRLGQA
GSGGGGTGDSEGSGALPSLTCSLTPLGLALVLWTVLGPC

Important features:

Signal peptide:
amino acids 1-26

Leucine zipper pattern.
amino acids 135-156

Glycosaminoglycan attachment site.
amino acids 436-439

N-glycosylation site.
amino acids 82-85, 179-183, 237-240, 372-375 and 423-426

VWFC domain
amino acids 411-425

FIGURE 162

GGAAGTCCACGGGGAGCTTGGATGCCAAAGGGAGGACGGCTGGGTCCTCTGGAGAGGACTAC
TCACTGGCATATTTCTGAGGTATCTGTAGAATAACCACAGCCTCAGATACTGGGGACTTTAC
AGTCCCACAGAACCGTCCTCCCAGGAAGCTGAATCCAGCAAGAACAATGGAGGCCAGCGGGA
AGCTCATTTGCAGACAAAGGCAAGTCCTTTTTTCCTTTCTCCTTTTGGGCTTATCTCTGGCG
GGCGCGGCGGAACCTAGAAGCTATTCTGTGGTGGAGGAAACTGAGGGCAGCTCCTTTGTCAC
CAATTTAGCAAAGGACCTGGGTCTGGAGCAGAGGGAATTCTCCAGGCGGGGGGTTAGGGTTG
TTTCCAGAGGGAACAAACTACATTTGCAGCTCAATCAGGAGACCGCGGATTTGTTGCTAAAT
GAGAAATTGGACCGTGAGGATCTGTGCGGTCACACAGAGCCCTGTGTGCTACGTTTCCAAGT
GTTGCTAGAGAGTCCCTTCGAGTTTTTTCAAGCTGAGCTGCAAGTAATAGACATAAACGACC
ACTCTCCAGTATTTCTGGACAAACAAATGTTGGTGAAAGTATCAGAGAGCAGTCCTCCTGGG
ACTACGTTTCCTCTGAAGAATGCCGAAGACTTAGATGTAGGCCAAAACAATATTGAGAACTA
TATAATCAGCCCCAACTCCTATTTTCGGGTCCTCACCCGCAAACGCAGTGATGGCAGGAAAT
ACCCAGAGCTGGTGCTGGACAAAGCGCTGGACCGAGAGGAAGAAGCTGAGCTCAGGTTAACA
CTCACAGCACTGGATGGTGGCTCTCCGCCCAGATCTGGCACTGCTCAGGTCTACATCGAAGT
CCTGGATGTCAACGATAATGCCCCTGAATTTGAGCAGCCTTTCTATAGAGTGCAGATCTCTG
AGGACAGTCCGGTAGGCTTCCTGGTTGTGAAGGTCTCTGCCACGGATGTAGACACAGGAGTC
AACGGAGAGATTTCCTATTCACTTTTCCAAGCTTCAGAAGAGATTGGCAAAACCTTTAAGAT
CAATCCCTTGACAGGAGAAATTGAACTAAAAAAACAACTCGATTTCGAAAAACTTCAGTCCT
ATGAAGTCAATATTGAGGCAAGAGATGCTGGAACCTTTTCTGGAAAATGCACCGTTCTGATT
CAAGTGATAGATGTGAACGACCATGCCCCAGAAGTTACCATGTCTGCATTTACCAGCCCAAT
ACCTGAGAACGCGCCTGAAACTGTGGTTGCACTTTTCAGTGTTTCAGATCTTGATTCAGGAG
AAAATGGGAAAATTAGTTGCTCCATTCAGGAGGATCTACCCTTCCTCCTGAAATCCGCGGAA
AACTTTTACACCCTACTAACGGAGAGACCACTAGACAGAGAAAGCAGAGCGGAATACAACAT
CACTATCACTGTCACTGACTTGGGGACCCCTATGCTGATAACACAGCTCAATATGACCGTGC
TGATCGCCGATGTCAATGACAACGCTCCCGCCTTCACCCAAACCTCCTACACCCTGTTCGTC
CGCGAGAACAACAGCCCCGCCCTGCACATCCGCAGCGTCAGCGCTACAGACAGAGACTCAGG
CACCAACGCCCAGGTCACCTACTCGCTGCTGCCGCCCCAGGACCCGCACCTGCCCCTCACAT
CCCTGGTCTCCATCAACGCGGACAACGGCCACCTGTTCGCCCTCAGGTCTCTGGACTACGAG
GCCCTGCAGGGGTTCCAGTTCCGCGTGGGCGCTTCAGACCACGGCTCCCCGGCGCTGAGCAG
CGAGGCGCTGGTGCGCGTGGTGGTGCTGGACGCCAACGACAACTCGCCCTTCGTGCTGTACC
CGCTGCAGAACGGCTCCGCGCCCTGCACCGAGCTGGTGCCCCGGGCGGCCGAGCCGGGCTAC
CTGGTGACCAAGGTGGTGGCGGTGGACGGCGACTCGGGCCAGAACGCCTGGCTGTCGTACCA
GCTGCTCAAGGCCACGGAGCTCGGTCTGTTCGGCGTGTGGGCGCACAATGGCGAGGTGCGCA
CCGCCAGGCTGCTGAGCGAGCGCGACGCGGCCAAGCACAGGCTGGTGGTGCTGGTCAAGGAC
AATGGCGAGCCTCCGCGCTCGGCCACCGCCACGCTGCACGTGCTCCTGGTGGACGGCTTCTC
CCAGCCCTACCTGCCTCTCCCGGAGGCGGCCCCGACCCAGGCCCAGGCCGACTTGCTCACCG
TCTACCTGGTGGTGGCGTTGGCCTCGGTGTCTTCGCTCTTCCTCTTTTCGGTGCTCCTGTTC
GTGGCGGTGCGGCTGTGTAGGAGGAGCAGGGCGGCCTCGGTGGGTCGCTGCTTGGTGCCCGA
GGGCCCCCTTCCAGGGCATCTTGTGGACATGAGCGGCACCAGGACCCTATCCCAGAGCTACC
AGTATGAGGTGTGTCTGGCAGGAGGCTCAGGGACCAATGAGTTCAAGTTCCTGAAGCCGATT
ATCCCCAACTTCCCTCCCCAGTGCCCTGGGAAAGAAATACAAGGAAATTCTACCTTCCCCAA
TAACTTTGGGTTCAATATTCAGTGACCATAGTTGACTTTTACATTCCATAGGTATTTTATTT
TGTGGCATTTCCATGCCAATGTTTATTTCCCCCAATTTGTGTGTATGTAATATTGTACGGAT
TTACTCTTGATTTTTCTCATGTTCTTTCTCCCTTTGTTTTAAAGTGAACATTTACCTTTATT
CCTGGTTCTT

FIGURE 163

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48314
<subunit 1 of 1, 798 aa, 1 stop
<MW: 87552, pI: 4.84, NX(S/T): 5
MEASGKLICRQRQVLFSFLLLGLSLAGAAEPRSYSVVEETEGSSFVTNLAKDLGLEQREFSR
RGVRVVSRGNKLHLQLNQETADLLLNEKLDREDLCGHTEPCVLRFQVLLESPFEFFQAELQV
IDINDHSPVFLDKQMLVKVSESSPPGTTFPLKNAEDLDVGQNNIENYIISPNSYFRVLTRKR
SDGRKYPELVLDKALDREEEAELRLTLTALDGGSPPRSGTAQVYIEVLDVNDNAPEFEQPFY
RVQISEDSPVGFLVVKVSATDVDTGVNGEISYSLFQASEEIGKTFKINPLTGEIELKKQLDF
EKLQSYEVNIEARDAGTFSGKCTVLIQVIDVNDHAPEVTMSAFTSPIPENAPETVVALFSVS
DLDSGENGKISCSIQEDLPFLLKSAENFYTLLTERPLDRESRAEYNITITVTDLGTPMLITQ
LNMTVLIADVNDNAPAFTQTSYTLFVRENNSPALHIRSVSATDRDSGTNAQVTYSLLPPQDP
HLPLTSLVSINADNGHLFALRSLDYEALQGFQFRVGASDHGSPALSSEALVRVVVLDANDNS
PFVLYPLQNGSAPCTELVPRAAEPGYLVTKVVAVDGDSGQNAWLSYQLLKATELGLFGVWAH
NGEVRTARLLSERDAAKHRLVVLVKDNGEPPRSATATLHVLLVDGFSQPYLPLPEAAPTQAQ
ADLLTVYLVVALASVSSLFLFSVLLFVAVRLCRRSRAASVGRCLVPEGPLPGHLVDMSGTRT
LSQSYQYEVCLAGGSGTNEFKFLKPIIPNFPPQCPGKEIQGNSTFPNNFGFNIQ

Important features:

Signal peptide:
amino acids 1-26

Transmembrane domain:
amino acids 685-712

Cadherins extracellular repeated domain signature.
amino acids 122-132, 231-241, 336-346, 439-449 and 549-559

ATP/GTP-binding site motif A (P-loop).
amino acids 285-292

N-glycosylation site.
amino acids 418-421, 436-439, 567-570 and 786-789

FIGURE 164

ACCCACGCGTCCGCCCACGCGTCCGCCCACGCGTCCGCCCACGCGTCCGCGCGTAGCCGTGC
GCCGATTGCCTCTCGGCCTGGGCAATGGTCCCGGCTGCCGGTCGACGACCGCCCCGCGTCAT
GCGGCTCCTCGGCTGGTGGCAAGTATTGCTGTGGGTGCTGGGACTTCCCGTCCGCGGCGTGG
AGGTTGCAGAGGAAAGTGGTCGCTTATGGTCAGAGGAGCAGCCTGCTCACCCTCTCCAGGTG
GGGGCTGTGTACCTGGGTGAGGAGGAGCTCCTGCATGACCCGATGGGCCAGGACAGGGCAGC
AGAAGAGGCCAATGCGGTGCTGGGGCTGGACACCCAAGGCGATCACATGGTGATGCTGTCTG
TGATTCCTGGGGAAGCTGAGGACAAAGTGAGTTCAGAGCCTAGCGGCGTCACCTGTGGTGCT
GGAGGAGCGGAGGACTCAAGGTGCAACGTCCGAGAGAGCCTTTTCTCTCTGGATGGCGCTGG
AGCACACTTCCCTGACAGAGAAGAGGAGTATTACACAGAGCCAGAAGTGGCGGAATCTGACG
CAGCCCCGACAGAGGACTCCAATAACACTGAAAGTCTGAAATCCCCAAAGGTGAACTGTGAG
GAGAGAAACATTACAGGATTAGAAAATTTCACTCTGAAAATTTTAAATATGTCACAGGACCT
TATGGATTTTCTGAACCCAAACGGTAGTGACTGTACTCTAGTCCTGTTTTACACCCCGTGGT
GCCGCTTTTCTGCCAGTTTGGCCCCTCACTTTAACTCTCTGCCCCGGGCATTTCCAGCTCTT
CACTTTTTGGCACTGGATGCATCTCAGCACAGCAGCCTTTCTACCAGGTTTGGCACCGTAGC
TGTTCCTAATATTTTATTATTTCAAGGAGCTAAACCAATGGCCAGATTTAATCATACAGATC
GAACACTGGAAACACTGAAAATCTTCATTTTTAATCAGACAGGTATAGAAGCCAAGAAGAAT
GTGGTGGTAACTCAAGCCGACCAAATAGGCCCTCTTCCCAGCACTTTGATAAAAAGTGTGGA
CTGGTTGCTTGTATTTTCCTTATTCTTTTTAATTAGTTTTATTATGTATGCTACCATTCGAA
CTGAGAGTATTCGGTGGCTAATTCCAGGACAAGAGCAGGAACATGTGGAGTAGTGATGGTCT
GAAAGAAGTTGGAAAGAGGAACTTCAATCCTTCGTTTCAGAAATTAGTGCTACAGTTTCATA
CATTTTCTCCAGTGACGTGTTGACTTGAAACTTCAGGCAGATTAAAAGAATCATTTGTTGAA
CAACTGAATGTATAAAAAAATTATAAACTGGTGTTTTAACTAGTATTGCAATAAGCAAATGC
AAAAATATTCAATAG

FIGURE 165

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48333
><subunit 1 of 1, 360 aa, 1 stop
><MW: 39885, pI: 4.79, NX(S/T): 7
MVPAAGRRPPRVMRLLGWWQVLLWVLGLPVRGVEVAEESGRLWSEEQPAHPLQVGAVYLGEE
ELLHDPMGQDRAAEEANAVLGLDTQGDHMVMLSVIPGEAEDKVSSEPSGVTCGAGGAEDSRC
NVRESLFSLDGAGAHFPDREEEYYTEPEVAESDAAPTEDSNNTESLKSPKVNCEERNITGLE
NFTLKILNMSQDLMDFLNPNGSDCTLVLFYTPWCRFSASLAPHFNSLPRAFPALHFLALDAS
QHSSLSTRFGTVAVPNILLFQGAKPMARFNHTDRTLETLKIFIFNQTGIEAKKNVVVTQADQ
IGPLPSTLIKSVDWLLVFSLFFLISFIMYATIRTESIRWLIPGQEQEHVE

Important features:

Signal peptide:
amino acids 1-25

Transmembrane domain:
amino acids 321-340

Homologous region to dilsufide isomerase
amino acids 212-302

N-glycosylation site.
amino acids 165-168, 181-184, 187-190, 194-197, 206-209, 278-281 and 293-296

Thioredoxin domain
amino acids 211-227

FIGURE 166

CCCGGCTCCGCTCCCTCTGCCCCCTCGGGGTCGCGCGCCCACGATGCTGCAGGGCCCTGGCT
CGCTGCTGCTGCTCTTCCTCGCCTCGCACTGCTGCCTGGGCTCGGCGCGCGGGCTCTTCCTC
TTTGGCCAGCCCGACTTCTCCTACAAGCGCAGCAATTGCAAGCCCATCCCGGTCAACCTGCA
GCTGTGCCACGGCATCGAATACCAGAACATGCGGCTGCCCAACCTGCTGGGCCACGAGACCA
TGAAGGAGGTGCTGGAGCAGGCCGGCGCTTGGATCCCGCTGGTCATGAAGCAGTGCCACCCG
GACACCAAGAAGTTCCTGTGCTCGCTCTTCGCCCCCGTCTGCCTCGATGACCTAGACGAGAC
CATCCAGCCATGCCACTCGCTCTGCGTGCAGGTGAAGGACCGCTGCGCCCCGGTCATGTCCG
CCTTCGGCTTCCCCTGGCCCGACATGCTTGAGTGCGACCGTTTCCCCCAGGACAACGACCTT
TGCATCCCCCTCGCTAGCAGCGACCACCTCCTGCCAGCCACCGAGGAAGCTCCAAAGGTATG
TGAAGCCTGCAAAAATAAAAATGATGATGACAACGACATAATGGAAACGCTTTGTAAAAATG
ATTTTGCACTGAAAATAAAAGTGAAGGAGATAACCTACATCAACCGAGATACCAAAATCATC
CTGGAGACCAAGAGCAAGACCATTTACAAGCTGAACGGTGTGTCCGAAAGGGACCTGAAGAA
ATCGGTGCTGTGGCTCAAAGACAGCTTGCAGTGCACCTGTGAGGAGATGAACGACATCAACG
CGCCCTATCTGGTCATGGGACAGAAACAGGGTGGGGAGCTGGTGATCACCTCGGTGAAGCGG
TGGCAGAAGGGGCAGAGAGAGTTCAAGCGCATCTCCCGCAGCATCCGCAAGCTGCAGTGCTA
GTCCCGGCATCCTGATGGCTCCGACAGGCCTGCTCCAGAGCACGGCTGACCATTTCTGCTCC
GGGATCTCAGCTCCCGTTCCCCAAGCACACTCCTAGCTGCTCCAGTCTCAGCCTGGGCAGCT
TCCCCCTGCCTTTTGCACGTTTGCATCCCCAGCATTTCCTGAGTTATAAGGCCACAGGAGTG
GATAGCTGTTTTCACCTAAAGGAAAAGCCCACCCGAATCTTGTAGAAATATTCAAACTAATA
AAATCATGAATATTTTAA

FIGURE 167

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50920

><subunit 1 of 1, 295 aa, 1 stop

><MW: 33518, pI: 7.74, NX(S/T): 0

MLQGPGSLLLLFLASHCCLGSARGLFLFGQPDFSYKRSNCKPIPVNLQLCHGIEYQNMRLPN
LLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAPVCLDDLDETIQPCHSLCVQVKDR
CAPVMSAFGFPWPDMLECDRFPQDNDLCIPLASSDHLLPATEEAPKVCEACKNKNDDDNDIM
ETLCKNDFALKIKVKEITYINRDTKIILETKSKTIYKLNGVSERDLKKSVLWLKDSLQCTCE
EMNDINAPYLVMGQKQGGELVITSVKRWQKGQREFKRISRSIRKLQC

Important features:

Signal peptide:

amino acids 1-20

Cysteine rich domain, homolgous to frizzled N terminus

amino acids 6-153

FIGURE 168

GTGGAGGCCGCCGACG**ATG**GCGGGGCCGACGGAGGCCGAGACGGGGTTGGCCGAGCCCCGGG
CCCTGTGCGCGCAGCGGGGCCACCGCACCTACGCGCGCCGCTGGGTGTTCCTGCTCGCGATC
AGCCTGCTCAACTGCTCCAACGCCACGCTGTGGCTCAGCTTTGCACCTGTGGCTGACGTCAT
TGCTGAGGACTTGGTCCTGTCCATGGAGCAGATCAACTGGCTGTCACTGGTCTACCTCGTGG
TATCCACCCCATTTGGCGTGGCGGCCATCTGGATCCTGGACTCCGTCGGGCTCCGTGCGGCG
ACCATCCTGGGTGCGTGGCTGAACTTTGCCGGGAGTGTGCTACGCATGGTGCCCTGCATGGT
TGTTGGGACCCAAAACCCATTTGCCTTCCTCATGGGTGGCCAGAGCCTCTGTGCCCTTGCCC
AGAGCCTGGTCATCTTCTCTCCAGCCAAGCTGGCTGCCTTGTGGTTCCCAGAGCACCAGCGA
GCCACGGCCAACATGCTCGCCACCATGTCGAACCCTCTGGGCGTCCTTGTGGCCAATGTGCT
GTCCCCTGTGCTGGTCAAGAAGGGTGAGGACATTCCGTTAATGCTCGGTGTCTATACCATCC
CTGCTGGCGTCGTCTGCCTGCTGTCCACCATCTGCCTGTGGGAGAGTGTGCCCCCCACCCCG
CCCTCTGCCGGGGCTGCCAGCTCCACCTCAGAGAAGTTCCTGGATGGGCTCAAGCTGCAGCT
CATGTGGAACAAGGCCTATGTCATCCTGGCTGTGTGCTTGGGGGGAATGATCGGGATCTCTG
CCAGCTTCTCAGCCCTCCTGGAGCAGATCCTCTGTGCAAGCGGCCACTCCAGTGGGTTTTCC
GGCCTCTGTGGCGCTCTCTTCATCACGTTTGGGATCCTGGGGGCACTGGCTCTCGGCCCCTA
TGTGGACCGGACCAAGCACTTCACTGAGGCCACCAAGATTGGCCTGTGCCTGTTCTCTCTGG
CCTGCGTGCCCTTTGCCCTGGTGTCCCAGCTGCAGGGACAGACCCTTGCCCTGGCTGCCACC
TGCTCGCTGCTCGGGCTGTTTGGCTTCTCGGTGGGCCCCGTGGCCATGGAGTTGGCGGTCGA
GTGTTCCTTCCCCGTGGGGGAGGGGGCTGCCACAGGCATGATCTTTGTGCTGGGGCAGGCCG
AGGGAATACTCATCATGCTGGCAATGACGGCACTGACTGTGCGACGCTCGGAGCCGTCCTTG
TCCACCTGCCAGCAGGGGGAGGATCCACTTGACTGGACAGTGTCTCTGCTGCTGATGGCCGG
CCTGTGCACCTTCTTCAGCTGCATCCTGGCGGTCTTCTTCCACACCCCATACCGGCGCCTGC
AGGCCGAGTCTGGGGAGCCCCCCTCCACCCGTAACGCCGTGGGCGGCGCAGACTCAGGGCCG
GGTGTGGACCGAGGGGGAGCAGGAAGGGCTGGGGTCCTGGGGCCCAGCACGGCGACTCCGGA
GTGCACGGCGAGGGGGGCCTCGCTAGAGGACCCCAGAGGGCCCGGGAGCCCCCACCCAGCCT
GCCACCGAGCGACTCCCCGTGCGCAAGGCCCAGCAGCCACCGACGCGCCCTCCCGCCCCGGC
AGACTCGCAGGCAGGGTCCAAGCGTCCAGGTTTATTGACCCGGCTGGGTCTCACTCCTCCTT
CTCCTCCCCGTGGGTGATCACG**TAG**CTGAGCGCCTTGTAGTCCAGGTTGCCCGCCACATCGA
TGGAGGCGAACTGGAACATCTGGTCCACCTGCGGGCGGGGGCGAAAGGGCTCCTTGCGGGCT
CCGGGAGCGAATTACAAGCGCGCACCTGAAAA

FIGURE 169

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50988
><subunit 1 of 1, 560 aa, 1 stop
><MW: 58427, pI: 6.86, NX(S/T): 2

MAGPTEAETGLAEPRALCAQRGHRTYARRWVFLLAISLLNCSNATLWLSFAPVADVIAEDLV
LSMEQINWLSLVYLVVSTPFGVAAIWILDSVGLRAATILGAWLNFAGSVLRMVPCMVVGTQN
PFAFLMGGQSLCALAQSLVIFSPAKLAALWFPEHQRATANMLATMSNPLGVLVANVLSPVLV
KKGEDIPLMLGVYTIPAGVVCLLSTICLWESVPPTPPSAGAASSTSEKFLDGLKLQLMWNKA
YVILAVCLGGMIGISASFSALLEQILCASGHSSGFSGLCGALFITFGILGALALGPYVDRTK
HFTEATKIGLCLFSLACVPFALVSQLQGQTLALAATCSLLGLFGFSVGPVAMELAVECSFPV
GEGAATGMIFVLGQAEGILIMLAMTALTVRRSEPSLSTCQQGEDPLDWTVSLLLMAGLCTFF
SCILAVFFHTPYRRLQAESGEPPSTRNAVGGADSGPGVDRGGAGRAGVLGPSTATPECTARG
ASLEDPRGPGSPHPACHRATPRAQGPAATDAPSRPGRLAGRVQASRFIDPAGSHSSFSSPWVIT

Important features:

Potential Transmembrane domains:

amino acids 30-50, 61-79, 98-112, 126-146, 169-182, 201-215, 248-268, 280-300, 318-337, 341-357, 375-387, 420-441

N-glycosylation site.

amino acids 40-43 and 43-46

Glycosaminoglycan attachment site.

amino acids 468-471

FIGURE 170A

GTCCCACATCCTGCTCAACTGGGTCAGGTCCCTCTTAGACCAGCTCTTGTCCATCATTTGCT
GAAGTGGACCAACTAGTTCCCCAGTAGGGGGTCTCCCCTGGCAATTCTTGATCGGCGTTTGG
ACATCTCAGATCGCTTCCAATGAAGATGGCCTTGCCTTGGGGTCCTGCTTGTTTCATAATCA
TCTAACTATGGGACAAGGTTGTGCCGGCAGCTCTGGGGGAAGGAGCACGGGGCTGATCAAGC
CATCCAGGAAACACTGGAGGACTTGTCCAGCCTTGAAAGAACTCTAGTGGTTTCTGAATCTA
GCCCACTTGGCGGTAAGC**ATG**ATGCAACTTCTGCAACTTCTGCTGGGGCTTTTGGGGCCAGG
TGGCTACTTATTTCTTTTAGGGGATTGTCAGGAGGTGACCACTCTCACGGTGAAATACCAAG
TGTCAGAGGAAGTGCCATCTGGTACAGTGATCGGGAAGCTGTCCCAGGAACTGGGCCGGGAG
GAGAGGCGGAGGCAAGCTGGGGCCGCCTTCCAGGTGTTGCAGCTGCCTCAGGCGCTCCCCAT
TCAGGTGGACTCTGAGGAAGGCTTGCTCAGCACAGGCAGGCGGCTGGATCGAGAGCAGCTGT
GCCGACAGTGGGATCCCTGCCTGGTTTCCTTTGATGTGCTTGCCACAGGGGATTTGGCTCTG
ATCCATGTGGAGATCCAAGTGCTGGACATCAATGACCACCAGCCACGGTTTCCCAAAGGCGA
GCAGGAGCTGGAAATCTCTGAGAGCGCCTCTCTGCGAACCCGGATCCCCCTGGACAGAGCTC
TTGACCCAGACACAGGCCCTAACACCCTGCACACCTACACTCTGTCTCCCAGTGAGCACTTT
GCCTTGGATGTCATTGTGGGCCCTGATGAGACCAAACATGCAGAACTCATAGTGGTGAAGGA
GCTGGACAGGGAAATCCATTCATTTTTTGATCTGGTGTTAACTGCCTATGACAATGGGAACC
CCCCCAAGTCAGGTACCAGCTTGGTCAAGGTCAACGTCTTGGACTCCAATGACAATAGCCCT
GCGTTTGCTGAGAGTTCACTGGCACTGGAAATCCAAGAAGATGCTGCACCTGGTACGCTTCT
CATAAAACTGACCGCCACAGACCCTGACCAAGGCCCCAATGGGGAGGTGGAGTTCTTCCTCA
GTAAGCACATGCCTCCAGAGGTGCTGGACACCTTCAGTATTGATGCCAAGACAGGCCAGGTC
ATTCTGCGTCGACCTCTAGACTATGAAAAGAACCCTGCCTACGAGGTGGATGTTCAGGCAAG
GGACCTGGGTCCCAATCCTATCCCAGCCCATTGCAAAGTTCTCATCAAGGTTCTGGATGTCA
ATGACAACATCCCAAGCATCCACGTCACATGGGCCTCCCAGCCATCACTGGTGTCAGAAGCT
CTTCCCAAGGACAGTTTTATTGCTCTTGTCATGGCAGATGACTTGGATTCAGGACACAATGG
TTTGGTCCACTGCTGGCTGAGCCAAGAGCTGGGCCACTTCAGGCTGAAAAGAACTAATGGCA
ACACATACATGTTGCTAACCAATGCCACACTGGACAGAGAGCAGTGGCCCAAATATACCCTC
ACTCTGTTAGCCCAAGACCAAGGACTCCAGCCCTTATCAGCCAAGAAACAGCTCAGCATTCA
GATCAGTGACATCAACGACAATGCACCTGTGTTTGAGAAAAGCAGGTATGAAGTCTCCACGC
GGGAAAACAACTTACCCTCTCTTCACCTCATTACCATCAAGGCTCATGATGCAGACTTGGGC
ATTAATGGAAAAGTCTCATACCGCATCCAGGACTCCCCAGTTGCTCACTTAGTAGCTATTGA
CTCCAACACAGGAGAGGTCACTGCTCAGAGGTCACTGAACTATGAAGAGATGGCCGGCTTTG
AGTTCCAGGTGATCGCAGAGGACAGCGGGCAACCCATGCTTGCATCCAGTGTCTCTGTGTGG
GTCAGCCTCTTGGATGCCAATGATAATGCCCCAGAGGTGGTCCAGCCTGTGCTCAGCGATGG
AAAAGCCAGCCTCTCCGTGCTTGTGAATGCCTCCACAGGCCACCTGCTGGTGCCCATCGAGA
CTCCCAATGGCTTGGGCCCAGCGGGCACTGACACACCTCCACTGGCCACTCACAGCTCCCGG
CCATTCCTTTTGACAACCATTGTGGCAAGAGATGCAGACTCGGGGGCAAATGGAGAGCCCCT
CTACAGCATCCGCAATGGAAATGAAGCCCACCTCTTCATCCTCAACCCTCATACGGGGCAGC
TGTTCGTCAATGTCACCAATGCCAGCAGCCTCATTGGGAGTGAGTGGGAGCTGGAGATAGTA
GTAGAGGACCAGGGAAGCCCCCCCTTACAGACCCGAGCCCTGTTGAGGGTCATGTTTGTCAC
CAGTGTGGACCACCTGAGGGACTCAGCCCGCAAGCCTGGGGCCTTGAGCATGTCGATGCTGA
CGGTGATCTGCCTGGCTGTACTGTTGGGCATCTTCGGGTTGATCCTGGCTTTGTTCATGTCC
ATCTGCCGGACAGAAAAGAAGGACAACAGGGCCTACAACTGTCGGGAGGCCGAGTCCACCTA
CCGCCAGCAGCCCAAGAGGCCCCAGAAACACATTCAGAAGGCAGACATCCACCTCGTGCCTG
TGCTCAGGGGTCAGGCAGGTGAGCCTTGTGAAGTCGGGCAGTCCCACAAAGATGTGGACAAG
GAGGCGATGATGGAAGCAGGCTGGGACCCCTGCCTGCAGGCCCCCTTCCACCTCACCCCGAC
CCTGTACAGGACGCTGCGTAATCAAGGCAACCAGGGAGCACCGGCGGAGAGCCGAGAGGTGC
TGCAAGACACGGTCAACCTCCTTTTCAACCATCCCAGGCAGAGGAATGCCTCCCGGGAGAAC
CTGAACCTTCCCGAGCCCCAGCCTGCCACAGGCCAGCCACGTTCCAGGCCTCTGAAGGTTGC
AGGCAGCCCCACAGGGAGGCTGGCTGGAGACCAGGGCAGTGAGGAAGCCCCACAGAGGCCAC
CAGCCTCCTCTGCAACCCTGAGACGGCAGCGACATCTCAATGGCAAAGTGTCCCCTGAGAAA
GAATCAGGGCCCCGTCAGATCCTGCGGAGCCTGGTCCGGCTGTCTGTGGCTGCCTTCGCCGA
GCGGAACCCCGTGGAGGAGCTCACTGTGGATTCTCCTCCTGTTCAGCAAATCTCCCAGCTGC
TGTCCTTGCTGCATCAGGGCCAATTCCAGCCCAAACCAAACCACCGAGGAAATAAGTACTTG
GCCAAGCCAGGAGGCAGCAGGAGTGCAATCCCAGACACAGATGGCCCAAGTGCAAGGGCTGG

FIGURE 170B

AGGCCAGACAGACCCAGAACAGGAGGAAGGGCCTTTGGATCCTGAAGAGGACCTCTCTGTGA
AGCAACTGCTAGAAGAAGAGCTGTCAAGTCTGCTGGACCCCAGCACAGGTCTGGCCCTGGAC
CGGCTGAGCGCCCCTGACCCGGCCTGGATGGCGAGACTCTCTTTGCCCCTCACCACCAACTA
CCGTGACAATGTGATCTCCCCGGATGCTGCAGCCACGGAGGAGCCGAGGACCTTCCAGACGT
TCGGCAAGGCAGAGGCACCAGAGCTGAGCCCAACAGGCACGAGGCTGGCCAGCACCTTTGTC
TCGGAGATGAGCTCACTGCTGGAGATGCTGCTGGAACAGCGCTCCAGCATGCCCGTGGAGGC
CGCCTCCGAGGCGCTGCGGCGGCTCTCGGTCTGCGGGAGGACCCTCAGTTTAGACTTGGCCA
CCAGTGCAGCCTCAGGCATGAAAGTGCAAGGGGACCCAGGTGGAAAGACGGGGACTGAGGGC
AAGAGCAGAGGCAGCAGCAGCAGCAGCAGGTGCCTGTGAACATACCTCAGACGCCTCTGGAT
CCAAGAACCAGGGGCCTGAGGATCTGTGGACAAGAGCTGGTTTCTAAAATCTTGTAACTCAC
TAGCTAGCGGCGGCCTGAGAACTTTAGGGTGACTGATGCTACCCCCACAGAGGAGGCAAGAG
CCCCAGGACTAACAGCTGACTGACCAAAGCAGCCCCTTGTAAGCAGCTCTGAGTCTTTTGGA
GGACAGGGACGGTTTGTGGCTGAGATAAGTGTTTCCTGGCAAAACATATGTGGAGCACAAAG
GGTCAGTCCTCTGGCAGAACAGATGCCACGGAGTATCACAGGCAGGAAAGGGTGGCCTTCTT
GGGTAGCAGGAGTCAGGGGGCTGTACCCTGGGGGTGCCAGGAAATGCTCTCTGACCTATCAA
TAAAGGAAAAGCAGTAAAAAAAAAAAAAAAAAAA

FIGURE 171

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48331
<subunit 1 of 1, 1184 aa, 1 stop
<MW: 129022, pI: 5.20, NX(S/T): 5
MMQLLQLLLGLLGPGGYLFLLGDCQEVTTLTVKYQVSEEVPSGTVIGKLSQELGREERRRQA
GAAFQVLQLPQALPIQVDSEEGLLSTGRRLDREQLCRQWDPCLVSFDVLATGDLALIHVEIQ
VLDINDHQPRFPKGEQELEISESASLRTRIPLDRALDPDTGPNTLHTYTLSPSEHFALDVIV
GPDETKHAELIVVKELDREIHSFFDLVLTAYDNGNPPKSGTSLVKVNVLDSNDNSPAFAESS
LALEIQEDAAPGTLLIKLTATDPDQGPNGEVEFFLSKHMPPEVLDTFSIDAKTGQVILRRPL
DYEKNPAYEVDVQARDLGPNPIPAHCKVLIKVLDVNDNIPSIHVTWASQPSLVSEALPKDSF
IALVMADDLDSGHNGLVHCWLSQELGHFRLKRTNGNTYMLLTNATLDREQWPKYTLTLLAQD
QGLQPLSAKKQLSIQISDINDNAPVFEKSRYEVSTRENNLPSLHLITIKAHDADLGINGKVS
YRIQDSPVAHLVAIDSNTGEVTAQRSLNYEEMAGFEFQVIAEDSGQPMLASSVSVWVSLLDA
NDNAPEVVQPVLSDGKASLSVLVNASTGHLLVPIETPNGLGPAGTDTPPLATHSSRPFLLTT
IVARDADSGANGEPLYSIRNGNEAHLFILNPHTGQLFVNVTNASSLIGSEWELEIVVEDQGS
PPLQTRALLRVMFVTSVDHLRDSARKPGALSMSMLTVICLAVLLGIFGLILALFMSICRTEK
KDNRAYNCREAESTYRQQPKRPQKHIQKADIHLVPVLRGQAGEPCEVGQSHKDVDKEAMMEA
GWDPCLQAPFHLTPTLYRTLRNQGNQGAPAESREVLQDTVNLLFNHPRQRNASRENLNLPEP
QPATGQPRSRPLKVAGSPTGRLAGDQGSEEAPQRPPASSATLRRQRHLNGKVSPEKESGPRQ
ILRSLVRLSVAAFAERNPVEELTVDSPPVQQISQLLSLLHQGQFQPKPNHRGNKYLAKPGGS
RSAIPDTDGPSARAGGQTDPEQEEGPLDPEEDLSVKQLLEEELSSLLDPSTGLALDRLSAPD
PAWMARLSLPLTTNYRDNVISPDAAATEEPRTFQTFGKAEAPELSPTGTRLASTFVSEMSSL
LEMLLEQRSSMPVEAASEALRRLSVCGRTLSLDLATSAASGMKVQGDPGGKTGTEGKSRGSS
SSSRCL

Important features:

Signal peptide:
amino acids 1-13

Transmembrane domain:
amino acids 719-739

N-glycosylation site.
amino acids 415-418, 582-585, 659-662, 662-665 amd 857-860

Cadherins extracellular repeated domain signature.
amino acids 123-133, 232-242, 340-350, 448-458 and 553-563

FIGURE 172

CGGACGCGTGGGCGGACGCGTGGGGGGAGAGCCGCAGTCCCGGCTGCAGCACCTGGGAGAAGG
CAGACCGTGTGAGGGGGCCTGTGGCCCCAGCGTGCTGTGGCCTCGGGGAGTGGGAAGTGGAG
GCAGGAGCCTTCCTTACACTTCGCCATGAGTTTCCTCATCGACTCCAGCATCATGATTACCT
CCCAGATACTATTTTTTGGATTTGGGTGGCTTTTCTTCATGCGCCAATTGTTTAAAGACTAT
GAGATACGTCAGTATGTTGTACAGGTGATCTTCTCCGTGACGTTTGCATTTTCTTGCACCAT
GTTTGAGCTCATCATCTTTGAAATCTTAGGAGTATTGAATAGCAGCTCCCGTTATTTTCACT
GGAAAATGAACCTGTGTGTAATTCTGCTGATCCTGGTTTTCATGGTGCCTTTTTACATTGGC
TATTTTATTGTGAGCAATATCCGACTACTGCATAAACAACGACTGCTTTTTTCCTGTCTCTT
ATGGCTGACCTTTATGTATTTCTTCTGGAAACTAGGAGATCCCTTTCCCATTCTCAGCCCAA
AACATGGGATCTTATCCATAGAACAGCTCATCAGCCGGGTTGGTGTGATTGGAGTGACTCTC
ATGGCTCTTCTTTCTGGATTTGGTGCTGTCAACTGCCCATACACTTACATGTCTTACTTCCT
CAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGCGACTGCTGCAAACCATGGATA
TGATCATAAGCAAAAAGAAAAGGATGGCAATGGCACGGAGAACAATGTTCCAGAAGGGGGAA
GTGCATAACAAACCATCAGGTTTCTGGGGAATGATAAAAAGTGTTACCACTTCAGCATCAGG
AAGTGAAAATCTTACTCTTATTCAACAGGAAGTGGATGCTTTGGAAGAATTAAGCAGGCAGC
TTTTTCTGGAAACAGCTGATCTATATGCTACCAAGGAGAGAATAGAATACTCCAAAACCTTC
AAGGGGAAATATTTTAATTTTCTTGGTTACTTTTTCTCTATTTACTGTGTTTGGAAAATTTT
CATGGCTACCATCAATATTGTTTTTGATCGAGTTGGGAAAACGGATCCTGTCACAAGAGGCA
TTGAGATCACTGTGAATTATCTGGGAATCCAATTTGATGTGAAGTTTTGGTCCCAACACATT
TCCTTCATTCTTGTTGGAATAATCATCGTCACATCCATCAGAGGATTGCTGATCACTCTTAC
CAAGTTCTTTTATGCCATCTCTAGCAGTAAGTCCTCCAATGTCATTGTCCTGCTATTAGCAC
AGATAATGGGCATGTACTTTGTCTCCTCTGTGCTGCTGATCCGAATGAGTATGCCTTTAGAA
TACCGCACCATAATCACTGAAGTCCTTGGAGAACTGCAGTTCAACTTCTATCACCGTTGGTT
TGATGTGATCTTCCTGGTCAGCGCTCTCTCTAGCATACTCTTCCTCTATTTGGCTCACAAAC
AGGCACCAGAGAAGCAAATGGCACCTTGAACTTAAGCCTACTACAGACTGTTAGAGGCCAGT
GGTTTCAAAATTTAGATATAAGAGGGGGGAAAAATGGAACCAGGGCCTGACATTTTATAAAC
AAACAAAATGCTATGGTAGCATTTTTCACCTTCATAGCATACTCCTTCCCCGTCAGGTGATA
CTATGACCATGAGTAGCATCAGCCAGAACATGAGAGGGAGAACTAACTCAAGACAATACTCA
GCAGAGAGCATCCCGTGTGGATATGAGGCTGGTGTAGAGGCGGAGAGGAGCCAAGAAACTAA
AGGTGAAAAATACACTGGAACTCTGGGGCAAGACATGTCTATGGTAGCTGAGCCAAACACGT
AGGATTTCCGTTTTAAGGTTCACATGGAAAAGGTTATAGCTTTGCCTTGAGATTGACTCATT
AAAATCAGAGACTGTAACAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCG
ACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATG

FIGURE 173

MSFLIDSSIMITSQILFFGFGWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEI
LGVLNSSSRYFHWKMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLTFMYFF
WKLGDPFPILSPKHGILSIEQLISRVGVIGVTLMALLSGFGAVNCPYTYMSYFLRNVTDTDI
LALERRLLQTMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWGMIKSVTTSASGSENLTLIQ
QEVDALEELSRQLFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVF
DRVGKTDPVTRGIEITVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISS
SKSSNVIVLLLAQIMGMYFVSSVLLIRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLVSA
LSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398, 425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 174

CATGGGAAGTGGAGCCGGAGCCTTCCTTACACTCGCCATGAGTTTCCTCATCGACTCCAGCA
TCATGATTACCTCCCNGANACTATTTTTTGGATTTGGGTGGCTTTTCTTCNGCGCCAATGTT
TAAAGACTATGAGATACGTCAGTATGTTGTACNGGTGATCTTCTCCGTGACGTTTGCCATTT
CTTGCACCATGTTTGAGCTCATCATCTTTGAAATCTTNGGAGTATTGAATAGCAGCTCCCGT
TATTTTCACTGGAAAATGAACCTGTGTGTAATTCTGCTGATCCTGGTTNTCATGGTGCCTTT
TTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCATAAACAACGACTGCTTTTTT
CCTGTCTCTTATGGCTGACCTTTATGTATTTCCAG

FIGURE 175

GTGTTGCCCTTGGGGAGGGGAAGGGGAGCCNGGCCCTTTCCTAAAATTTGGCCAAGGGTTTC
TTTNTTGAATTCCGGGTTNNGNATACCTTCCCAGAAAATATTTTTTGGATTTGGGGTAGNTT
TTTTTCATGCGCCAATTGTTTAAAGACTATGAGATACGTCAGTATGTTGTACAGGTGATNTT
NTCCGTGACGTTTGCATTTTCTTGCACCATGTTTGAGCTCATCATNTTTGAAATNTTAGGAG
TATTGAATAGCAGCTCCCGTTATTTTCACTGGAAAATGAACCTGTGTGTAATTCTGCTGATC
CTGGTTTTCATGGTGCCTTTTTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCA
TAAACAACGACTGCTTTTTTCCTGTCTNTTATGGCTGACCTTTATGTATTTNTTNTGGAAAN
TAGGAGATCCCTTTCCCATTCTC

FIGURE 176A

CTCGCGCAGGGATCGTCCCATGGCCGGGGCTCGGAGCCGCGACCCTTGGGGGGCCTCCGGGA
TTTGCTACCTTTTTGGCTCCCTGCTCGTCGAACTGCTCTTCTCACGGGCTGTCGCCTTCAAT
CTGGACGTGATGGGTGCCTTGCGCAAGGAGGGCGAGCCAGGCAGCCTCTTCGGCTTCTCTGT
GGCCCTGCACCGGCAGTTGCAGCCCCGACCCCAGAGCTGGCTGCTGGTGGGTGCTCCCCAGG
CCCTGGCTCTTCCTGGGCAGCAGGCGAATCGCACTGGAGGCCTCTTCGCTTGCCCGTTGAGC
CTGGAGGAGACTGACTGCTACAGAGTGGACATCGACCAGGGAGCTGATATGCAAAAGGAAAG
CAAGGAGAACCAGTGGTTGGGAGTCAGTGTTCGGAGCCAGGGGCCTGGGGGCAAGATTGTTA
CCTGTGCACACCGATATGAGGCAAGGCAGCGAGTGGACCAGATCCTGGAGACGCGGGATATG
ATTGGTCGCTGCTTTGTGCTCAGCCAGGACCTGGCCATCCGGGATGAGTTGGATGGTGGGGA
ATGGAAGTTCTGTGAGGGACGCCCCCAAGGCCATGAACAATTTGGGTTCTGCCAGCAGGGCA
CAGCTGCCGCCTTCTCCCCTGATAGCCACTACCTCCTCTTTGGGGCCCCAGGAACCTATAAT
TGGAAGGGCACGGCCAGGGTGGAGCTCTGTGCACAGGGCTCAGCGGACCTGGCACACCTGGA
CGACGGTCCCTACGAGGCGGGGGGAGAGAAGGAGCAGGACCCCCGCCTCATCCCGGTCCCTG
CCAACAGCTACTTTGGCTTCTCTATTGACTCGGGGAAAGGTCTGGTGCGTGCAGAAGAGCTG
AGCTTTGTGGCTGGAGCCCCCCGCGCCAACCACAAGGGTGCTGTGGTCATCCTGCGCAAGGA
CAGCGCCAGTCGCCTGGTGCCCGAGGTTATGCTGTCTGGGGAGCGCCTGACCTCCGGCTTTG
GCTACTCACTGGCTGTGGCTGACCTCAACAGTGATGGCTGGCCAGACCTGATAGTGGGTGCC
CCCTACTTCTTTGAGCGCCAAGAAGAGCTGGGGGGTGCTGTGTATGTGTACTTGAACCAGGG
GGGTCACTGGGCTGGGATCTCCCCTCTCCGGCTCTGCGGCTCCCCTGACTCCATGTTCGGGA
TCAGCCTGGCTGTCCTGGGGGACCTCAACCAAGATGGCTTTCCAGATATTGCAGTGGGTGCC
CCCTTTGATGGTGATGGGAAAGTCTTCATCTACCATGGGAGCAGCCTGGGGGTTGTCGCCAA
ACCTTCACAGGTGCTGGAGGGCGAGGCTGTGGGCATCAAGAGCTTCGGCTACTCCCTGTCAG
GCAGCTTGGATATGGATGGGAACCAATACCCTGACCTGCTGGTGGGCTCCCTGGCTGACACC
GCAGTGCTCTTCAGGGCCAGACCCATCCTCCATGTCTCCCATGAGGTCTCTATTGCTCCACG
AAGCATCGACCTGGAGCAGCCCAACTGTGCTGGCGGCCACTCGGTCTGTGTGGACCTAAGGG
TCTGTTTCAGCTACATTGCAGTCCCCAGCAGCTATAGCCCTACTGTGGCCCTGGACTATGTG
TTAGATGCGGACACAGACCGGAGGCTCCGGGGCCAGGTTCCCCGTGTGACGTTCCTGAGCCG
TAACCTGGAAGAACCCAAGCACCAGGCCTCGGGCACCGTGTGGCTGAAGCACCAGCATGACC
GAGTCTGTGGAGACGCCATGTTCCAGCTCCAGGAAAATGTCAAAGACAAGCTTCGGGCCATT
GTAGTGACCTTGTCCTACAGTCTCCAGACCCCTCGGCTCCGGCGACAGGCTCCTGGCCAGGG
GCTGCCTCCAGTGGCCCCCATCCTCAATGCCCACCAGCCCAGCACCCAGCGGGCAGAGATCC
ACTTCCTGAAGCAAGGCTGTGGTGAAGACAAGATCTGCCAGAGCAATCTGCAGCTGGTCCAC
GCCCGCTTCTGTACCCGGGTCAGCGACACGGAATTCCAACCTCTGCCCATGGATGTGGATGG
AACAACAGCCCTGTTTGCACTGAGTGGGCAGCCAGTCATTGGCCTGGAGCTGATGGTCACCA
ACCTGCCATCGGACCCAGCCCAGCCCCAGGCTGATGGGGATGATGCCCATGAAGCCCAGCTC
CTGGTCATGCTTCCTGACTCACTGCACTACTCAGGGGTCCGGGCCCTGGACCCTGCGGAGAA
GCCACTCTGCCTGTCCAATGAGAATGCCTCCCATGTTGAGTGTGAGCTGGGGAACCCCATGA
AGAGAGGTGCCCAGGTCACCTTCTACCTCATCCTTAGCACCTCCGGGATCAGCATTGAGACC
ACGGAACTGGAGGTAGAGCTGCTGTTGGCCACGATCAGTGAGCAGGAGCTGCATCCAGTCTC
TGCACGAGCCCGTGTCTTCATTGAGCTGCCACTGTCCATTGCAGGAATGGCCATTCCCCAGC
AACTCTTCTTCTCTGGTGTGGTGAGGGGCGAGAGAGCCATGCAGTCTGAGCGGGATGTGGGC
AGCAAGGTCAAGTATGAGGTCACGGTTTCCAACCAAGGCCAGTCGCTCAGAACCCTGGGCTC
TGCCTTCCTCAACATCATGTGGCCTCATGAGATTGCCAATGGGAAGTGGTTGCTGTACCCAA
TGCAGGTTGAGCTGGAGGGCGGGCAGGGGCCTGGGCAGAAAGGGCTTTGCTCTCCCAGGCCC
AACATCCTCCACCTGGATGTGGACAGTAGGGATAGGAGGCGGCGGGAGCTGGAGCCACCTGA
GCAGCAGGAGCCTGGTGAGCGGCAGGAGCCCAGCATGTCCTGGTGGCCAGTGTCCTCTGCTG
AGAAGAAGAAAAACATCACCCTGGACTGCGCCCGGGGCACGGCCAACTGTGTGGTGTTCAGC
TGCCCACTCTACAGCTTTGACCGCGCGGCTGTGCTGCATGTCTGGGGCCGTCTCTGGAACAG
CACCTTTCTGGAGGAGTACTCAGCTGTGAAGTCCCTGGAAGTGATTGTCCGGGCCAACATCA
CAGTGAAGTCCTCCATAAAGAACTTGATGCTCCGAGATGCCTCCACAGTGATCCCAGTGATG
GTATACTTGGACCCCATGGCTGTGGTGGCAGAAGGAGTGCCCTGGTGGGTCATCCTCCTGGC
TGTACTGGCTGGGCTGCTGGTGCTAGCACTGCTGGTGCTGCTCCTGTGGAAGATGGGATTCT
TCAAACGGGCGAAGCACCCCGAGGCCACCGTGCCCCAGTACCATGCGGTGAAGATTCCTCGG
GAAGACCGACAGCAGTTCAAGGAGGAGAAGACGGGCACCATCCTGAGGAACAACTGGGGCAG

FIGURE 176B

CCCCCGGCGGGAGGGCCCGGATGCACACCCCATCCTGGCTGCTGACGGGCATCCCGAGCTGG
GCCCCGATGGGCATCCAGGGCCAGGCACCGCCTAGGTTCCCATGTCCCAGCCTGGCCTGTGG
CTGCCCTCCATCCCTTCCCCAGAGATGGCTCCTTGGGATGAAGAGGGTAGAGTGGGCTGCTG
GTGTCGCATCAAGATTTGGCAGGATCGGCTTCCTCAGGGGCACAGACCTCTCCCACCCACAA
GAACTCCTCCCACCCAACTTCCCCTTAGAGTGCTGTGAGATGAGAGTGGGTAAATCAGGGAC
AGGGCCATGGGGTAGGGTGAGAAGGGCAGGGGTGTCCTGATGCAAAGGTGGGGAGAAGGGAT
CCTAATCCCTTCCTCTCCCATTCACCCTGTGTAACAGGACCCCAAGGACCTGCCTCCCCGGA
AGTGCCTTAACCTAGAGGGTCGGGGAGGAGGTTGTGTCACTGACTCAGGCTGCTCCTTCTCT
AGTTTCCCCTCTCATCTGACCTTAGTTTGCTGCCATCAGTCTAGTGGTTTCGTGGTTTCGTC
TATTTATTAAAAAATATTTGAGAACAAAAAAAAAAAAAAAAAAA

FIGURE 177

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA55737

><subunit 1 of 1, 1141 aa, 1 stop

><MW: 124671, pI: 5.82, NX(S/T): 5

MAGARSRDPWGASGICYLFGSLLVELLFSRAVAFNLDVMGALRKEGEPGSLFGFSVALHRQL
QPRPQSWLLVGAPQALALPGQQANRTGGLFACPLSLEETDCYRVDIDQGADMQKESKENQWL
GVSVRSQGPGGKIVTCAHRYEARQRVDQILETRDMIGRCFVLSQDLAIRDELDGGEWKFCEG
RPQGHEQFGFCQQGTAAAFSPDSHYLLFGAPGTYNWKGTARVELCAQGSADLAHLDDGPYEA
GGEKEQDPRLIPVPANSYFGFSIDSGKGLVRAEELSFVAGAPRANHKGAVVILRKDSASRLV
PEVMLSGERLTSGFGYSLAVADLNSDGWPDLIVGAPYFFERQEELGGAVYVYLNQGGHWAGI
SPLRLCGSPDSMFGISLAVLGDLNQDGFPDIAVGAPFDGDGKVFIYHGSSLGVVAKPSQVLE
GEAVGIKSFGYSLSGSLDMDGNQYPDLLVGSLADTAVLFRARPILHVSHEVSIAPRSIDLEQ
PNCAGGHSVCVDLRVCFSYIAVPSSYSPTVALDYVLDADTDRRLRGQVPRVTFLSRNLEEPK
HQASGTVWLKHQHDRVCGDAMFQLQENVKDKLRAIVVTLSYSLQTPRLRRQAPGQGLPPVAP
ILNAHQPSTQRAEIHFLKQGCGEDKICQSNLQLVHARFCTRVSDTEFQPLPMDVDGTTALFA
LSGQPVIGLELMVTNLPSDPAQPQADGDDAHEAQLLVMLPDSLHYSGVRALDPAEKPLCLSN
ENASHVECELGNPMKRGAQVTFYLILSTSGISIETTELEVELLLATISEQELHPVSARARVF
IELPLSIAGMAIPQQLFFSGVVRGERAMQSERDVGSKVKYEVTVSNQGQSLRTLGSAFLNIM
WPHEIANGKWLLYPMQVELEGGQGPGQKGLCSPRPNILHLDVDSRDRRRRELEPPEQQEPGE
RQEPSMSWWPVSSAEKKKNITLDCARGTANCVVFSCPLYSFDRAAVLHVWGRLWNSTFLEEY
SAVKSLEVIVRANITVKSSIKNLMLRDASTVIPVMVYLDPMAVVAEGVPWWVILLAVLAGLL
VLALLVLLLWKMGFFKRAKHPEATVPQYHAVKIPREDRQQFKEEKTGTILRNNWGSPRREGP
DAHPILAADGHPELGPDGHPGPGTA

Important features:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 1039-1064

N-glycosylation sites.

amino acids 86-89, 746-749, 949-952, 985-988 and 1005-1008

Integrins alpha chain proteins.

amino acids 1064-1071, 384-408, 1041-1071, 317-346, 443-465, 385-407, 215-224, 634-647, 85-99, 322-346, 470-479, 442-466, 379-408 and 1031-1047

FIGURE 178

CGCGCCGGGCGCAGGGAGCTGAGTGGACGGCTCGAGACGGCGGCGCGTGCAGCAGCTCCAGA
AAGCAGCGAGTTGGCAGAGCAGGGCTGCATTTCCAGCAGGAGCTGCGAGCACAGTGCTGGCT
CACAACAAGATGCTCAAGGTGTCAGCCGTACTGTGTGTGTGCAGCCGCTTGGTGCAGTCA
GTCTCTCGCAGCTGCCGCGGCGGTGGCTGCAGCCGGGGGGCGGTCGGACGGCGGTAATTTTC
TGGATGATAAACAATGGCTCACCACAATCTCTCAGTATGACAAGGAAGTCGGACAGTGGAAC
AAATTCCGAGACGAAGTAGAGGATGATTATTTCCGCACTTGGAGTCCAGGAAAACCCTTCGA
TCAGGCTTTAGATCCAGCTAAGGATCCATGCTTAAAGATGAAATGTAGTCGCCATAAAGTAT
GCATTGCTCAAGATTCTCAGACTGCAGTCTGCATTAGTCACCGGAGGCTTACACACAGGATG
AAAGAAGCAGGAGTAGACCATAGGCAGTGGAGGGGTCCCATATTATCCACCTGCAAGCAGTG
CCCAGTGGTCTATCCCAGCCCTGTTTGTGGTTCAGATGGTCATACCTACTCTTTTCAGTGCA
AACTAGAATATCAGGCATGTGTCTTAGGAAAACAGATCTCAGTCAAATGTGAAGGACATTGC
CCATGTCCTTCAGATAAGCCCACCAGTACAAGCAGAAATGTTAAGAGAGCATGCAGTGACCT
GGAGTTCAGGGAAGTGGCAAACAGATTGCGGGACTGGTTCAAGGCCCTTCATGAAAGTGGAA
GTCAAAACAAGAAGACAAAAACATTGCTGAGGCCTGAGAGAAGCAGATTCGATACCAGCATC
TTGCCAATTTGCAAGGACTCACTTGGCTGGATGTTTAACAGACTTGATACAAACTATGACCT
GCTATTGGACCAGTCAGAGCTCAGAAGCATTTACCTTGATAAGAATGAACAGTGTACCAAGG
CATTCTTCAATTCTTGTGACACATACAAGGACAGTTTAATATCTAATAATGAGTGGTGCTAC
TGCTTCCAGAGACAGCAAGACCCACCTTGCCAGACTGAGCTCAGCAATATTCAGAAGCGGCA
AGGGGTAAAGAAGCTCCTAGGACAGTATATCCCCCTGTGTGATGAAGATGGTTACTACAAGC
CAACACAATGTCATGGCAGTGTTGGACAGTGCTGGTGTGTTGACAGATATGGAAATGAAGTC
ATGGGATCCAGAATAAATGGTGTTGCAGATTGTGCTATAGATTTTGAGATCTCCGGAGATTT
TGCTAGTGGCGATTTTCATGAATGGACTGATGATGAGGATGATGAAGACGATATTATGAATG
ATGAAGATGAAATTGAAGATGATGATGAAGATGAAGGGGATGATGATGATGGTGGTGATGAC
CATGATGTATACATTTGATTGATGACAGTTGAAATCAATAAATTCTACATTTCTAATATTTA
CAAAAATGATAGCCTATTTAAAATTATCTTCTTCCCCAATAACAAAATGATTCTAAACCTCA
CATATATTTTGTATAATTATTTGAAAAATTGCAGCTAAAGTTATAGAACTTTATGTTTAAAT
AAGAATCATTTGCTTTGAGTTTTTATATTCCTTACACAAAAAGAAAATACATATGCAGTCTA
GTCAGACAAAATAAAGTTTTGAAGTGCTACTATAATAAATTTTTCACGAGAACAAACTTTGT
AAATCTTCCATAAGCAAAATGACAGCTAGTGCTTGGGATCGTACATGTTAATTTTTTGAAAG
ATAATTCTAAGTGAAATTTAAAATAAATAAATTTTTAATGACCTGGGTCTTAAGGATTTAGG
AAAAATATGCATGCTTTAATTGCATTTCCAAAGTAGCATCTTGCTAGACCTAGATGAGTCAG
GATAACAGAGAGATACCACATGACTCCAAAAAAAAAAAAAAA

FIGURE 179

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49829
><subunit 1 of 1, 436 aa, 1 stop
><MW: 49429, pI: 4.80, NX(S/T): 0

MLKVSAVLCVCAAAWCSQSLAAAAAVAAAGGRSDGGNFLDDKQWLTTISQYDKEVGQWNKFR
DEVEDDYFRTWSPGKPFDQALDPAKDPCLKMKCSRHKVCIAQDSQTAVCISHRRLTHRMKEA
GVDHRQWRGPILSTCKQCPVVYPSPVCGSDGHTYSFQCKLEYQACVLGKQISVKCEGHCPCP
SDKPTSTSRNVKRACSDLEFREVANRLRDWFKALHESGSQNKKTKTLLRPERSRFDTSILPI
CKDSLGWMFNRLDTNYDLLLDQSELRSIYLDKNEQCTKAFFNSCDTYKDSLISNNEWCYCFQ
RQQDPPCQTELSNIQKRQGVKKLLGQYIPLCDEDGYYKPTQCHGSVGQCWCVDRYGNEVMGS
RINGVADCAIDFEISGDFASGDFHEWTDDEDDEDDIMNDEDEIEDDDEDEGDDDDGGDDHDVYI

Important features:

Signal peptide:
amino acids 1-16

Leucine zipper pattern.
amino acids 246-267

N-myristoylation sites.
amino acids 357-362, 371-376 and 376-381

Thyroglobulin type-1 repeat proteins
amino acids 353-365 and 339-352

FIGURE 180A

CAGACTCCAGATTTCCCTGTCAACCACGAGGAGTCCAGAGAGGAAACGCGGAGCGGAGACAA
CAGTACCTGACGCCTCTTTCAGCCCGGGATCGCCCCAGCAGGGATGGGCGACAAGATCTGGC
TGCCCTTCCCCGTGCTCCTTCTGGCCGCTCTGCCTCCGGTGCTGCTGCCTGGGGCGGCCGGC
TTCACACCTTCCCTCGATAGCGACTTCACCTTTACCCTTCCCGCCGGCCAGAAGGAGTGCTT
CTACCAGCCCATGCCCCTGAAGGCCTCGCTGGAGATCGAGTACCAAGTTTTAGATGGAGCAG
GATTAGATATTGATTTCCATCTTGCCTCTCCAGAAGGCAAAACCTTAGTTTTTGAACAAAGA
AAATCAGATGGAGTTCACACTGTAGAGACTGAAGTTGGTGATTACATGTTCTGCTTTGACAA
TACATTCAGCACCATTTCTGAGAAGGTGATTTTCTTTGAATTAATCCTGGATAATATGGGAG
AACAGGCACAAGAACAAGAAGATTGGAAGAAATATATTACTGGCACAGATATATTGGATATG
AAACTGGAAGACATCCTGGAATCCATCAACAGCATCAAGTCCAGACTAAGCAAAAGTGGGCA
CATACAAATTCTGCTTAGAGCATTTGAAGCTCGTGATCGAAACATACAAGAAAGCAACTTTG
ATAGAGTCAATTTCTGGTCTATGGTTAATTTAGTGGTCATGGTGGTGGTGTCAGCCATTCAA
GTTTATATGCTGAAGAGTCTGTTTGAAGATAAGAGGAAAAGTAGAACTTAAAACTCCAAACT
AGAGTACGTAACATTGAAAAATGAGGCATAAAAATGCAATAAACTGTTACAGTCAAGACCAT
TAATGGTCTTCTCCAAAATATTTTGAGATATAAAAGTAGGAAACAGGTATAATTTTAATGTG
AAAATTAAGTCTTCACTTTCTGTGCAAGTAATCCTGCTGATCCAGTTGTACTTAAGTGTGTA
ACAGGAATATTTTGCAGAATATAGGTTTAACTGAATGAAGCCATATTAATAACTGCATTTTC
CTAACTTTGAAAAATTTTGCAAATGTCTTAGGTGATTTAAATAAATGAGTATTGGGCCTAAT
TGCAACACCAGTCTGTTTTTAACAGGTTCTATTACCCAGAACTTTTTTGTAAATGCGGCAGT
TACAAATTAACTGTGGAAGTTTTCAGTTTTAAGTTATAAATCACCTGAGAATTACCTAATGA
TGGATTGAATAAATCTTTAGACTACAAAAGCCCAACTTTTCTCTATTTACATATGCATCTCT
CCTATAATGTAAATAGAATAATAGCTTTGAAATACAATTAGGTTTTTGAGATTTTTATAACC
AAATACATTTCAGTGTAACATATTAGCAGAAAGCATTAGTCTTTGTACTTTGCTTACATTCC
CAAAAGCTGACATTTTCACGATTCTTAAAAACACAAAGTTACACTTACTAAAATTAGGACAT
GTTTTCTCTTTGAAATGAAGAATATAGTTTAAAAGCTTCCTCCTCCATAGGGACACATTTTC
TCTAACCCTTAACTAAAGTGTAGGATTTTAAAATTAAATGTGAGGTAAAATAAGTTTATTTT
TAATAGTATCTGTCAAGTTAATATCTGTCAACAGTTAATAATCATGTTATGTTAATTTTAAC
ATGATTGCTGACTTGGATAATTCATTATTACCAGCAGTTATGAAGGAAATATTGCTAAAATG
ATCTGGGCCTACCATAAATAAATATCTCCTTTTCTGAGCTCTAAGAATTATCAGAAAACAGG
AAAGAATTTAGAAAAACTTGAGAAAACCTAATCCAAAATAAAATTCACTTAAGTAGAACTAT
AAATAAATATCTAGAATCTGACTGGCTCATCATGACATCCTACTCATAACATAAATCAAAGG
AGATGATTAATTTCCAGTTAGCTGGAAGAAACTTTGGCTGTAGGTTTTTATTTTCTACAAGA
ATTCTGGTTTGAATTATTTTTGTAAGCAGGTACATTTTATAAAATGTAAGCCCTACTGTAAG
GTTTAGCACTGGGTGTACATATTTATTAAAAATTTTTATTATAACAACTTTTATTAAAATGG
CCTTTCTGAACACTTTATTTATTGATGTTGAAGTAAGGATTAGAAACATAGACTCCCAAGTT
TTAAACACCTAAATGTGAATAACCCATATATACAACAAAGTTTCTGCCATCTAGCTTTTTGA
AGTCTATGGGGGTCTTACTCAAGTACTAGTAATTTAACTTCATCATGAATGAACTATAATTT
TTAAGTTATGCCCATTTATAACGTTGTTTATGACTACATTGTGAGTTAGAAACAAACTTAAA
ATTTGGGGTATAGAACCCCTCAACAGGTTAGTAATGCTGGAATTCTTGATGAGCAATAATGA
TAACCAGAGAGTGATTTCATTTACACTCATAGTAGTATAAAAAGAGATACATTTCCCTCTTA
GGCCCCTGGGAGAAGAGCAGCTTAGATTTCCCTACTGGCAAGGTTTTTAAAAATGAGGTAAA
TGCCGTATATGATCAATTACCTTAATTGGCCAAGAAAATGCTTCAGGTGTCTAGGGGTATCC
TCTGCAACACTTGCAGAACAAAGGTCAATAAGATCCTTGCCTATGAATACCCCTCCCTTTTG
CGCTGTTAAATTTGCAATGAGAAGCAAATTTACAGTACCATAACTAATAAAGCAGGGTACAG
ATATAAACTACTGCATCTTTTCTATAAAACTGTGATTAAGAATTCTACCTCTCCTGTATGGC
TGTTACTGTACTGTACTCTCTGACTCCTTACCTAACAATGAATTTGTTACATAATCTTCTAC
ATGTATGATTTGTGCCACTGATCTTAAACCTATGATTCAGTAACTTCTTACCATATAAAAAC
GATAATTGCTTTATTTGGAAAAGAATTTAGGAATACTAAGGACAATTATTTTTATAGACAAA
GTAAAAAGACAGATATTTAAGAGGCATAACCAAAAAAGCAAAACTTGTAAACAGAGTAAAAA
TCTTTAATATTTCTAAAGACATACTGTTTATCTGCTTCATATGCTTTTTTTAATTTCACTAT
TCCATTTCTAAATTAAAGTTATGCTAAATTGAGTAAGCTGTTTATCACTTAACAGCTCATTT
TGTCTTTTTCAATATACAAATTTTAAAAATACTACAATATTTAACTAAGGCCCAACCGATTT
CCATAATGTAGCAGTTACCGTGTTCACCTCACACTAAGGCCTAGAGTTTGCTCTGATATGCA
TTTGGATGATTAATGTTATGCTGTTCTTTCATGTGAATGTCAAGACATGGAGGGTGTTTGTA

FIGURE 180B

ATTTTATGGTAAAATTAATCCTTCTTACACATAATGGTGTCTTAAAATTGACAAAAAATGAG
CACTTACAATTGTATGTCTCCTCAAATGAAGATTCTTTATGTGAAATTTTAAAAGACATTGA
TTCCGCATGTAAGGATTTTTCATCTGAAGTACAATAATGCACAATCAGTGTTGCTCAAACTG
CTTTATACTTATAAACAGCCATCTTAAATAAGCAACGTATTGTGAGTACTGATATGTATATA
ATAAAAATTATCAAAGGAAAA

FIGURE 181

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52196
><subunit 1 of 1, 229 aa, 1 stop
><MW: 26017, pI: 4.73, NX(S/T): 0
MGDKIWLPFPVLLLAALPPVLLPGAAGFTPSLDSDFTFTLPAGQKECFYQPMPLKASLEIEY
QVLDGAGLDIDFHLASPEGKTLVFEQRKSDGVHTVETEVGDYMFCFDNTFSTISEKVIFFEL
ILDNMGEQAQEQEDWKKYITGTDILDMKLEDILESINSIKSRLSKSGHIQILLRAFEARDRN
IQESNFDRVNFWSMVNLVVMVVVSAIQVYMLKSLFEDKRKSRT

Important features:

Signal peptide:
amino acids 1-23

Transmembrane domain:
amino acids 195-217

N-myristoylation site.
amino acids 43-48

Tyrosine kinase phosphorylation site.
amino acids 55-62

FIGURE 182

CCATCCCTGAGATCTTTTTATAAAAAACCCAGTCTTTGCTGACCAGACAAAGCATACCAGAT
CTCACCAGAGAGTCGCAGACACTATGCTGCCTCCCATGGCCCTGCCCAGTGTGTCCTGGATG
CTGCTTTCCTGCCTCATTCTCCTGTGTCAGGTTCAAGGTGAAGAAACCCAGAAGGAACTGCC
CTCTCCACGGATCAGCTGTCCCAAAGGCTCCAAGGCCTATGGCTCCCCCTGCTATGCCTTGT
TTTTGTCACCAAAATCCTGGATGGATGCAGATCTGGCTTGCCAGAAGCGGCCCTCTGGAAAA
CTGGTGTCTGTGCTCAGTGGGGCTGAGGGATCCTTCGTGTCCTCCCTGGTGAGGAGCATTAG
TAACAGCTACTCATACATCTGGATTGGGCTCCATGACCCCACACAGGGCTCTGAGCCTGATG
GAGATGGATGGGAGTGGAGTAGCACTGATGTGATGAATTACTTTGCATGGGAGAAAAATCCC
TCCACCATCTTAAACCCTGGCCACTGTGGGAGCCTGTCAAGAAGCACAGGATTTCTGAAGTG
GAAAGATTATAACTGTGATGCAAAGTTACCCTATGTCTGCAAGTTCAAGGACTAGGGCAGGT
GGGAAGTCAGCAGCCTCAGCTTGGCGTGCAGCTCATCATGGACATGAGACCAGTGTGAAGAC
TCACCCTGGAAGAGAATATTCTCCCCAAACTGCCCTACCTGACTACCTTGTCATGATCCTCC
TTCTTTTTCCTTTTTCTTCACCTTCATTTCAGGCTTTTCTCTGTCTTCCATGTCTTGAGATC
TCAGAGAATAATAATAAAAATGTTACTTTATAAAAAAAAAAAAAAAAAAAAAA

FIGURE 183

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56965

<subunit 1 of 1, 175 aa, 1 stop

<MW: 19330, pI: 7.25, NX(S/T): 1

MLPPMALPSVSWMLLSCLILLCQVQGEETQKELPSPRISCPKGSKAYGSPCYALFLSPKSWM
DADLACQKRPSGKLVSVLSGAEGSFVSSLVRSISNSYSYIWIGLHDPTQGSEPDGDGWEWSS
TDVMNYFAWEKNPSTILNPGHCGSLSRSTGFLKWKDYNCDAKLPYVCKFKD

Important features:

Signal peptide:

amino acids 1-26

C-type lectin domain signature.

amino acids 146-171

FIGURE 184

CCAGTCTGTCGCCACCTCACTTGGTGTCTGCTGTCCCCGCCAGGCAAGCCTGGGGTGAGAGC
ACAGAGGAGTGGGCCGGGACCATGCGGGGGACGCGGCTGGCGCTCCTGGCGCTGGTGCTGGC
TGCCTGCGGAGAGCTGGCGCCGGCCCTGCGCTGCTACGTCTGTCCGGAGCCCACAGGAGTGT
CGGACTGTGTCACCATCGCCACCTGCACCACCAACGAAACCATGTGCAAGACCACACTCTAC
TCCCGGGAGATAGTGTACCCCTTCCAGGGGGACTCCACGGTGACCAAGTCCTGTGCCAGCAA
GTGTAAGCCCTCGGATGTGGATGGCATCGGCCAGACCCTGCCCGTGTCCTGCTGCAATACTG
AGCTGTGCAATGTAGACGGGGCGCCCGCTCTGAACAGCCTCCACTGCGGGGCCCTCACGCTC
CTCCCACTCTTGAGCCTCCGACTGTAGAGTCCCCGCCCACCCCCATGGCCCTATGCGGCCCA
GCCCCGAATGCCTTGAAGAAGTGCCCCCTGCACCAGGAAAAAAAAAAAAAAAAA

FIGURE 185

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56405

<subunit 1 of 1, 125 aa, 1 stop

<MW: 13115, pI: 5.90, NX(S/T): 1

MRGTRLALLALVLAACGELAPALRCYVCPEPTGVSDCVTIATCTTNETMCKTTLYSREIVYP

FQGDSTVTKSCASKCKPSDVDGIGQTLPVSCCNTELCNVDGAPALNSLHCGALTLLPLLSLRL

Important features:

Signal peptide:

amino acids 1-17

N-glycosylation site.

amino acids 46-49

FIGURE 186

CTGCAGTCAGGACTCTGGGACCGCAGGGGGCTCCCGGACCCTGACTCTGCAGCCGAACCGGC
ACGGTTTCGTGGGGACCCAGGCTTGCAAAGTGACGGTCATTTTCTCTTTCTTTCTCCCTCTT
GAGTCCTTCTGAGATGATGGCTCTGGGCGCAGCGGGAGCTACCCGGGTCTTTGTCGCGATGG
TAGCGGCGGCTCTCGGCGGCCACCCTCTGCTGGGAGTGAGCGCCACCTTGAACTCGGTTCTC
AATTCCAACGCTATCAAGAACCTGCCCCCACCGCTGGGCGGCGCTGCGGGGCACCCAGGCTC
TGCAGTCAGCGCCGCGCCGGGAATCCTGTACCCGGGCGGGAATAAGTACCAGACCATTGACA
ACTACCAGCCGTACCCGTGCGCAGAGGACGAGGAGTGCGGCACTGATGAGTACTGCGCTAGT
CCCACCCGCGGAGGGGACGCAGGCGTGCAAATCTGTCTCGCCTGCAGGAAGCGCCGAAAACG
CTGCATGCGTCACGCTATGTGCTGCCCCGGGAATTACTGCAAAAATGGAATATGTGTGTCTT
CTGATCAAAATCATTTCCGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTTGGTAATGAT
CATAGCACCTTGGATGGGTATTCCAGAAGAACCACCTTGTCTTCAAAAATGTATCACACCAA
AGGACAAGAAGGTTCTGTTTGTCTCCGGTCATCAGACTGTGCCTCAGGATTGTGTTGTGCTA
GACACTTCTGGTCCAAGATCTGTAAACCTGTCCTGAAAGAAGGTCAAGTGTGTACCAAGCAT
AGGAGAAAAGGCTCTCATGGACTAGAAATATTCCAGCGTTGTTACTGTGGAGAAGGTCTGTC
TTGCCGGATACAGAAAGATCACCATCAAGCCAGTAATTCTTCTAGGCTTCACACTTGTCAGA
GACACTAAACCAGCTATCCAAATGCAGTGAACTCCTTTTATATAATAGATGCTATGAAAACC
TTTTATGACCTTCATCAACTCAATCCTAAGGATATACAAGTTCTGTGGTTTCAGTTAAGCAT
TCCAATAACACCTTCCAAAAACCTGGAGTGTAAGAGCTTTGTTTCTTTATGGAACTCCCCTG
TGATTGCAGTAAATTACTGTATTGTAAATTCTCAGTGTGGCACTTACCTGTAAATGCAATGA
AACTTTTAATTATTTTTCTAAAGGTGCTGCACTGCCTATTTTTCCTCTTGTTATGTAAATTT
TTGTACACATTGATTGTTATCTTGACTGACAAATATTCTATATTGAACTGAAGTAAATCATT
TCAGCTTATAGTTCTTAAAAGCATAACCCTTTACCCCATTTAATTCTAGAGTCTAGAACGCA
AGGATCTCTTGGAATGACAAATGATAGGTACCTAAAATGTAACATGAAAATACTAGCTTATT
TTCTGAAATGTACTATCTTAATGCTTAAATTATATTTCCCTTTAGGCTGTGATAGTTTTTGA
AATAAAATTTAACATTTAAAAAAAAAAAA

FIGURE 187

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57530

<subunit 1 of 1, 266 aa, 1 stop

<MW: 28672, pI: 8.85, NX(S/T): 1

MMALGAAGATRVFVAMVAAALGGHPLLGVSATLNSVLNSNAIKNLPPPLGGAAGHPGSAVSA
APGILYPGGNKYQTIDNYQPYPCAEDEECGTDEYCASPTRGGDAGVQICLACRKRRKRCMRH
AMCCPGNYCKNGICVSSDQNHFRGEIEETITESFGNDHSTLDGYSRRTTLSSKMYHTKGQEG
SVCLRSSDCASGLCCARHFWSKICKPVLKEGQVCTKHRRKGSHGLEIFQRCYCGEGLSCRIQ
KDHHQASNSSRLHTCQRH

Important features:

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 256-259

Fungal Zn(2)-Cys(6) binuclear cluster domain

amino acids 110-126

FIGURE 188

TGTGTTTCCCTGCAGTCAGAATTTGGGACNGCAGGGGTTCCCGGACCTGATTTTGCAGCGGA
ACGGGAAGGTTTTGTGGGACCCAGGTTGAAATGACGGTCATTTTTTTTTCTTTCTCCTTCNG
GAGTCCTTNTGAGANGATGGTTTTGGGCGCAGCGGGAGCTAACCCGGTTTTTTGTNGCGATG
GTAGCGGCGGTTTTCGGCGGCCACCTTNTGCTGGGAGTGAGCGCCACCTTGAATCGGTTTTC
AATTCCAACGNTATCAAGAACCTGCCCCCACCGNTGGGCGGCGCTGCGGGGCACCCAGGNTT
TGCAGTCAGCGCCGCGCCGGGAATCCTGTACCCGGGCGGGAATAAGTACCAGACCATTGACA
ATTACCAGCCGTACCCGTGCGCAGAGGACGAGGAGTGCGGCACTGATGAGTACTGCGCTAGT
CCCACCCGCGGAGGGGANGCGGGCGTGCAAATNTGTNTNGCCTGCAGGAAGCGCCGAAAACG
CTGCATGCGTCANGCTATGTGCTGCCCCGGGAATTACTGCAAAAATGGAATATGTGTGTNTT
CTGATCAAAATCATTTCCGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTTGGTAATGAT
CATAGCACCTTGGATGGG

FIGURE 189A

GAGGAACCTACCGGTACCGGCCGCGCGCTGGTAGTCGCCGGTGTGGCTGCACCTCACCAATC
CCGTGCGCCGCGGCTGGGCCGTCGGAGAGTGCGTGTGCTTCTCTCCTGCACGCGGTGCTTGG
GCTCGGCCAGGCGGGGTCCGCCGCCAGGGTTTGAGGATGGGGGAGTAGCTACAGGAAGCGAC
CCCGCGATGGCAAGGTATATTTTTGTGGAATGAAAAGGAAGTATTAGAAATGAGCTGAAGAC
CATTCACAGATTAATATTTTTGGGGACAGATTTGTGATGCTTGATTCACCCTTGAAGTAATG
TAGACAGAAGTTCTCAAATTTGCATATTACATCAACTGGAACCAGCAGTGAATCTTAATGTT
CACTTAAATCAGAACTTGCATAAGAAAGAGAATGGGAGTCTGGTTAAATAAAGATGACTATA
TCAGAGACTTGAAAAGGATCATTCTCTGTTTTCTGATAGTGTATATGGCCATTTTAGTGGGC
ACAGATCAGGATTTTTACAGTTTACTTGGAGTGTCCAAAACTGCAAGCAGTAGAGAAATAAG
ACAAGCTTTCAAGAAATTGGCATTGAAGTTACATCCTGATAAAAACCCGAATAACCCAAATG
CACATGGCGATTTTTTAAAAATAAATAGAGCATATGAAGTACTCAAAGATGAAGATCTACGG
AAAAAGTATGACAAATATGGAGAAAAGGGACTTGAGGATAATCAAGGTGGCCAGTATGAAAG
CTGGAACTATTATCGTTATGATTTTGGTATTTATGATGATGATCCTGAAATCATAACATTGG
AAAGAAGAGAATTTGATGCTGCTGTTAATTCTGGAGAACTGTGGTTTGTAAATTTTTACTCC
CCAGGCTGTTCACACTGCCATGATTTAGCTCCCACATGGAGAGACTTTGCTAAAGAAGTGGA
TGGGTTACTTCGAATTGGAGCTGTTAACTGTGGTGATGATAGAATGCTTTGCCGAATGAAAG
GAGTCAACAGCTATCCCAGTCTCTTCATTTTTCGGTCTGGAATGGCCCCAGTGAAATATCAT
GGAGACAGATCAAAGGAGAGTTTAGTGAGTTTTGCAATGCAGCATGTTAGAAGTACAGTGAC
AGAACTTTGGACAGGAAATTTTGTCAACTCCATACAAACTGCTTTTGCTGCTGGTATTGGCT
GGCTGATCACTTTTTTGTTCAAAAGGAGGAGATTGTTTGACTTCACAGACACGACTCAGGCTT
AGTGGCATGTTGTTTCTCAACTCATTGGATGCTAAAGAAATATATTTGGAAGTAATACATAA
TCTTCCAGATTTTGAACTACTTTCGGCAAACACACTAGAGGATCGTTTGGCTCATCATCGGT
GGCTGTTATTTTTTCATTTTGGAAAAAATGAAAATTCAAATGATCCTGAGCTGAAAAAACTA
AAAACTCTACTTAAAAATGATCATATTCAAGTTGGCAGGTTTGACTGTTCCTCTGCACCAGA
CATCTGTAGTAATCTGTATGTTTTTCAGCCGTCTCTAGCAGTATTTAAAGGACAAGGAACCA
AAGAATATGAAATTCATCATGGAAAGAAGATTCTATATGATATACTTGCCTTTGCCAAAGAA
AGTGTGAATTCTCATGTTACCACGCTTGGACCTCAAAATTTTCCTGCCAATGACAAAGAACC
ATGGCTTGTTGATTTCTTTGCCCCCTGGTGTCCACCATGTCGAGCTTTACTACCAGAGTTAC
GAAGAGCATCAAATCTTCTTTATGGTCAGCTTAAGTTTGGTACACTAGATTGTACAGTTCAT
GAGGGACTCTGTAACATGTATAACATTCAGGCTTATCCAACAACAGTGGTATTCAACCAGTC
CAACATTCATGAGTATGAAGGACATCACTCTGCTGAACAAATCTTGGAGTTCATAGAGGATC
TTATGAATCCTTCAGTGGTCTCCCTTACACCCACCACCTTCAACGAACTAGTTACACAAAGA
AAACACAACGAAGTCTGGATGGTTGATTTCTATTCTCCGTGGTGTCATCCTTGCCAAGTCTT
AATGCCAGAATGGAAAAGAATGGCCCGGACATTAACTGGACTGATCAACGTGGGCAGTATAG
ATTGCCAACAGTATCATTCTTTTTGTGCCCAGGAAAACGTTCAAAGATACCCTGAGATAAGA
TTTTTTCCCCCAAAATCAAATAAAGCTTATCAGTATCACAGTTACAATGGTTGGAATAGGGA
TGCTTATTCCCTGAGAATCTGGGGTCTAGGATTTTTACCTCAAGTATCCACAGATCTAACAC
CTCAGACTTTCAGTGAAAAAGTTCTACAAGGGAAAAATCATTGGGTGATTGATTTCTATGCT
CCTTGGTGTGGACCTTGCCAGAATTTTGCTCCAGAATTTGAGCTCTTGGCTAGGATGATTAA
AGGAAAAGTGAAAGCTGGAAAAGTAGACTGTCAGGCTTATGCTCAGACATGCCAGAAAGCTG
GGATCAGGGCCTATCCAACTGTTAAGTTTTATTTCTACGAAAGAGCAAAGAGAAATTTTCAA
GAAGAGCAGATAAATACCAGAGATGCAAAAGCAATCGCTGCCTTAATAAGTGAAAAATTGGA
AACTCTCCGAAATCAAGGCAAGAGGAATAAGGATGAACTTTGATAATGTTGAAGATGAAGAA
AAAGTTTAAAAGAAATTCTGACAGATGACATCAGAAGACACCTATTTAGAATGTTACATTTA
TGATGGGAATGAATGAACATTATCTTAGACTTGCAGTTGTACTGCCAGAATTATCTACAGCA
CTGGTGTAAAAGAAGGGTCTGCAAACTTTTTCTGTAAAGGGCCGGTTTATAAATATTTTAGA
CTTTGCAGGCTATAATATATGGTTCACACATGAGAACAAGAATAGAGTCATCATGTATTCTT
TGTTATTTGCTTTTAACAACCTTTAAAAAATATTAAAACGATTCTTAGCTCAGAGCCATACA
AAAGTAGGCTGGATTCAGTCCATGGACCATAGATTGCTGTCCCCCTCGACGGACTTATAATG
TTTCAGGTGGCTGGCTTGAACATGAGTCTGCTGTGCTATCTACATAAATGTCTAAGTTGTAT
AAAGTCCACTTTCCCTTCACGTTTTTTGGCTGACCTGAAAAGAGGTAACTTAGTTTTTGGTC
ACTTGTTCTCCTAAAAATGCTATCCCTAACCATATATTTATATTTCGTTTTAAAAACACCCA
TGATGTGGCACAGTAAACAAACCCTGTTATGCTGTATTATTATGAGGAGATTCTTCATTGTT
TTCTTTCCTTCTCAAAGGTTGAAAAAATGCTTTTAATTTTTCACAGCCGAGAAACAGTGCAG

FIGURE 189B

CAGTATATGTGCACACAGTAAGTACACAAATTTGAGCAACAGTAAGTGCACAAATTCTGTAG
TTTGCTGTATCATCCAGGAAAACCTGAGGGAAAAAAATTATAGCAATTAACTGGGCATTGTA
GAGTATCCTAAATATGTTATCAAGTATTTAGAGTTCTATATTTTAAAGATATATGTGTTCAT
GTATTTTCTGAAATTGCTTTCATAGAAATTTTCCCACTGATAGTTGATTTTTGAGGCATCTA
ATATTTACATATTTGCCTTCTGAACTTTGTTTTGACCTGTATCCTTTATTTACATTGGGTTT
TTCTTTCATAGTTTTGGTTTTTCACTCCTGTCCAGTCTATTTATTATTCAAATAGGAAAAAT
TACTTTACAGGTTGTTTTACTGTAGCTTATAATGATACTGTAGTTATTCCAGTTACTAGTTT
ACTGTCAGAGGGCTGCCTTTTTCAGATAAATATTGACATAATAACTGAAGTTATTTTTATAA
GAAAATCAAGTATATAAATCTAGGAAAGGGATCTTCTAGTTTCTGTGTTGTTTAGACTCAAA
GAATCACAAATTTGTCAGTAACATGTAGTTGTTTAGTTATAATTCAGAGTGTACAGAATGGT
AAAAATTCCAATCAGTCAAAAGAGGTCAATGAATTAAAAGGCTTGCAACTTTTTCAAAAAAA
AAAAAAAAAA

FIGURE 190

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56439
<subunit 1 of 1, 747 aa, 1 stop
<MW: 86127, pI: 7.46, NX(S/T): 2
MGVWLNKDDYIRDLKRIILCFLIVYMAILVGTDQDFYSLLGVSKTASSREIRQAFKKLALKL
HPDKNPNNPNAHGDFLKINRAYEVLKDEDLRKKYDKYGEKGLEDNQGGQYESWNYYRYDFGI
YDDDPEIITLERREFDAAVNSGELWFVNFYSPGCSHCHDLAPTWRDFAKEVDGLLRIGAVNC
GDDRMLCRMKGVNSYPSLFIFRSGMAPVKYHGDRSKESLVSFAMQHVRSTVTELWTGNFVNS
IQTAFAAGIGWLITFCSKGGDCLTSQTRLRLSGMLFLNSLDAKEIYLEVIHNLPDFELLSAN
TLEDRLAHHRWLLFFHFGKNENSNDPELKKLKTLLKNDHIQVGRFDCSSAPDICSNLYVFQP
SLAVFKGQGTKEYEIHHGKKILYDILAFAKESVNSHVTTLGPQNFPANDKEPWLVDFFAPWC
PPCRALLPELRRASNLLYGQLKFGTLDCTVHEGLCNMYNIQAYPTTVVFNQSNIHEYEGHHS
AEQILEFIEDLMNPSVVSLTPTTFNELVTQRKHNEVWMVDFYSPWCHPCQVLMPEWKRMART
LTGLINVGSIDCQQYHSFCAQENVQRYPEIRFFPPKSNKAYQYHSYNGWNRDAYSLRIWGLG
FLPQVSTDLTPQTFSEKVLQGKNHWVIDFYAPWCGPCQNFAPEFELLARMIKGKVKAGKVDC
QAYAQTCQKAGIRAYPTVKFYFYERAKRNFQEEQINTRDAKAIAALISEKLETLRNQGKRNKDEL

Important features:

Endoplasmic reticulum targeting sequence.
amino acids 744-747

Cytochrome c family heme-binding site signature.
amino acids 158-163

Nt-dnaJ domain signature.
amino acids 77-96

N-glycosylation site.
amino acids 484-487

FIGURE 191

AGACAGTACCTCCTCCCTAGGACTACACAAGGACTGAACCAGAAGGAAGAGGACAGAGCAAA
GCCATGAACATCATCCTAGAAATCCTTCTGCTTCTGATCACCATCATCTACTCCTACTTGGA
GTCGTTGGTGAAGTTTTTCATTCCTCAGAGGAGAAAATCTGTGGCTGGGGAGATTGTTCTCA
TTACTGGAGCTGGGCATGGAATAGGCAGGCAGACTACTTATGAATTTGCAAAACGACAGAGC
ATATTGGTTCTGTGGGATATTAATAAGCGCGGTGTGGAGGAAACTGCAGCTGAGTGCCGAAA
ACTAGGCGTCACTGCGCATGCGTATGTGGTAGACTGCAGCAACAGAGAAGAGATCTATCGCT
CTCTAAATCAGGTGAAGAAAGAAGTGGGTGATGTAACAATCGTGGTGAATAATGCTGGGACA
GTATATCCAGCCGATCTTCTCAGCACCAAGGATGAAGAGATTACCAAGACATTTGAGGTCAA
CATCCTAGGACATTTTTGGATCACAAAAGCACTTCTTCCATCGATGATGGAGAGAAATCATG
GCCACATCGTCACAGTGGCTTCAGTGTGCGGCCACGAAGGGATTCCTTACCTCATCCCATAT
TGTTCCAGCAAATTTGCCGCTGTTGGCTTTCACAGAGGTCTGACATCAGAACTTCAGGCCTT
GGGAAAAACTGGTATCAAAACCTCATGTCTCTGCCCAGTTTTTGTGAATACTGGGTTCACCA
AAAATCCAAGCACAAGATTATGGCCTGTATTGGAGACAGATGAAGTCGTAAGAAGTCTGATA
GATGGAATACTTACCAATAAGAAAATGATTTTTGTTCCATCGTATATCAATATCTTTCTGAG
ACTACAGAAGTTTCTTCCTGAACGCGCCTCAGCGATTTTAAATCGTATGCAGAATATTCAAT
TTGAAGCAGTGGTTGGCCACAAAATCAAAATGAAATGAATAAATAAGCTCCAGCCAGAGATG
TATGCATGATAATGATATGAATAGTTTCGAATCAATGCTGCAAAGCTTTATTTCACATTTTT
TCAGTCCTGATAATATTAAAAACATTGGTTTGGCACTAGCAGCAGTCAAACGAACAAGATTA
ATTACCTGTCTTCCTGTTTCTCAAGAATATTTACGTAGTTTTTCATAGGTCTGTTTTTCCTT
TCATGCCTCTTAAAAACTTCTGTGCTTACATAAACATACTTAAAAGGTTTTCTTTAAGATAT
TTTATTTTTCCATTTAAAGGTGGACAAAAGCTACCTCCCTAAAAGTAAATACAAAGAGAACT
TATTTACACAGGGAAGGTTTAAGACTGTTCAAGTAGCATTCCAATCTGTAGCCATGCCACAG
AATATCAACAAGAACACAGAATGAGTGCACAGCTAAGAGATCAAGTTTCAGCAGGCAGCTTT
ATCTCAACCTGGACATATTTTAAGATTCAGCATTTGAAAGATTTCCCTAGCCTCTTCCTTTT
TCATTAGCCCAAAACGGTGCAACTCTATTCTGGACTTTATTACTTGATTCTGTCTTCTGTAT
AACTCTGAAGTCCACCAAAAGTGGACCCTCTATATTTCCTCCCTTTTTATAGTCTTATAAGA
TACATTATGAAAGGTGACCGACTCTATTTTAAATCTCAGAATTTTAAGTTCTAGCCCCATGA
TAACCTTTTTCTTTGTAATTTATGCTTTCATATATCCTTGGTCCCAGAGATGTTTAGACAAT
TTTAGGCTCAAAAATTAAAGCTAACACAGGAAAAGGAACTGTACTGGCTATTACATAAGAAA
CAATGGACCCAAGAGAAGAA

FIGURE 192

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56409

<subunit 1 of 1, 300 aa, 1 stop

<MW: 33655, pI: 9.31, NX(S/T): 1

MNIILEILLLLITIIYSYLESLVKFFIPQRRKSVAGEIVLITGAGHGIGRQTTYEFAKRQSI
LVLWDINKRGVEETAAECRKLGVTAHAYVVDCSNREEIYRSLNQVKKEVGDVTIVVNNAGTV
YPADLLSTKDEEITKTFEVNILGHFWITKALLPSMMERNHGHIVTVASVCGHEGIPYLIPYC
SSKFAAVGFHRGLTSELQALGKTGIKTSCLCPVFVNTGFTKNPSTRLWPVLETDEVVRSLID
GILTNKKMIFVPSYINIFLRLQKFLPERASAILNRMQNIQFEAVVGHKIKMK

Important features:

Signal peptide:

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-33 and 58-61

Short-chain alcohol dehydrogenase family protein

amino acids 165-202, 37-49, 112-122 and 210-219

FIGURE 193

CGGCGGCGGCTGCGGGCGCGAGGTGAGGGGCGCGAGGTGAGGGGCGCGAGGTTCCCAGCAGG
ATGCCCCGGCTCTGCAGGAAGCTGAAGTGAGAGGCCCGGAGAGGGCCCAGCCCGCCCGGGGC
AGGATGACCAAGGCCCGGCTGTTCCGGCTGTGGCTGGTGCTGGGGTCGGTGTTCATGATCCT
GCTGATCATCGTGTACTGGGACAGCGCAGGCGCCGCGCACTTCTACTTGCACACGTCCTTCT
CTAGGCCGCACACGGGGCCGCCGCTGCCCACGCCCGGGCCGGACAGGGACAGGGAGCTCACG
GCCGACTCCGATGTCGACGAGTTTCTGGACAAGTTTCTCAGTGCTGGCGTGAAGCAGAGCGA
CCTTCCCAGAAAGGAGACGGAGCAGCCGCCTGCGCCGGGGAGCATGGAGGAGAGCGTGAGAG
GCTACGACTGGTCCCCGCGCGACGCCCGGCGCAGCCCAGACCAGGGCCGGCAGCAGGCGGAG
CGGAGGAGCGTGCTGCGGGGCTTCTGCGCCAACTCCAGCCTGGCCTTCCCCACCAAGGAGCG
CGCATTCGACGACATCCCCAACTCGGAGCTGAGCCACCTGATCGTGGACGACCGGCACGGGG
CCATCTACTGCTACGTGCCCAAGGTGGCCTGCACCAACTGGAAGCGCGTGATGATCGTGCTG
AGCGGAAGCCTGCTGCACCGCGGTGCGCCCTACCGCGACCCGCTGCGCATCCCGCGCGAGCA
CGTGCACAACGCCAGCGCGCACCTGACCTTCAACAAGTTCTGGCGCCGCTACGGGAAGCTCT
CCCGCCACCTCATGAAGGTCAAGCTCAAGAAGTACACCAAGTTCCTCTTCGTGCGCGACCCC
TTCGTGCGCCTGATCTCCGCCTTCCGCAGCAAGTTCGAGCTGGAGAACGAGGAGTTCTACCG
CAAGTTCGCCGTGCCCATGCTGCGGCTGTACGCCAACCACACCAGCCTGCCCGCCTCGGCGC
GCGAGGCCTTCCGCGCTGGCCTCAAGGTGTCCTTCGCCAACTTCATCCAGTACCTGCTGGAC
CCGCACACGGAGAAGCTGGCGCCCTTCAACGAGCACTGGCGGCAGGTGTACCGCCTCTGCCA
CCCGTGCCAGATCGACTACGACTTCGTGGGGAAGCTGGAGACTCTGGACGAGGACGCCGCGC
AGCTGCTGCAGCTACTCCAGGTGGACCGGCAGCTCCGCTTCCCCCCGAGCTACCGGAACAGG
ACCGCCAGCAGCTGGGAGGAGGACTGGTTCGCCAAGATCCCCCTGGCCTGGAGGCAGCAGCT
GTATAAACTCTACGAGGCCGACTTTGTTCTCTTCGGCTACCCCAAGCCCGAAAACCTCCTCC
GAGACTGAAAGCTTTCGCGTTGCTTTTTCTCGCGTGCCTGGAACCTGACGCACGCGCACTCC
AGTTTTTTTATGACCTACGATTTTGCAATCTGGGCTTCTTGTTCACTCCACTGCCTCTATCC
ATTGAGTACTGTATCGATATTGTTTTTTAAGATTAATATATTTCAGGTATTTAATACGA

FIGURE 194

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56112

<subunit 1 of 1, 414 aa, 1 stop

<MW: 48414, pI: 9.54, NX(S/T): 4

MTKARLFRLWLVLGSVFMILLIIVYWDSAGAAHFYLHTSFSRPHTGPPLPTPGPDRDRELTA
DSDVDEFLDKFLSAGVKQSDLPRKETEQPPAPGSMEESVRGYDWSPRDARRSPDQGRQQAER
RSVLRGFCANSSLAFPTKERAFDDIPNSELSHLIVDDRHGAIYCYVPKVACTNWKRVMIVLS
GSLLHRGAPYRDPLRIPREHVHNASAHLTFNKFWRRYGKLSRHLMKVKLKKYTKFLFVRDPF
VRLISAFRSKFELENEEFYRKFAVPMLRLYANHTSLPASAREAFRAGLKVSFANFIQYLLDP
HTEKLAPFNEHWRQVYRLCHPCQIDYDFVGKLETLDEDAAQLLQLLQVDRQLRFPPSYRNRT
ASSWEEDWFAKIPLAWRQQLYKLYEADFVLFGYPKPENLLRD

Important features:

Signal peptide:

amino acids 1-31

N-glycosylation sites.

amino acids 134-137, 209-212, 280-283 and 370-373

TNFR/NGFR family cysteine-rich region protein

amino acids 329-332

FIGURE 195

TCGGGCCAGAATTCGGCACGAGGCGGCACGAGGGCGACGGCCTCACGGGGCTTTGGAGGTGA
AAGAGGCCCAGAGTAGAGAGAGAGAGAGACCGACGTACACGGGATGGCTACGGGAACGCGCT
ATGCCGGGAAGGTGGTGGTCGTGACCGGGGGCGGGCGCGGCATCGGAGCTGGGATCGTGCGC
GCCTTCGTGAACAGCGGGGCCCGAGTGGTTATCTGCGACAAGGATGAGTCTGGGGGCCGGGC
CCTGGAGCAGGAGCTCCCTGGAGCTGTCTTTATCCTCTGTGATGTGACTCAGGAAGATGATG
TGAAGACCCTGGTTTCTGAGACCATCCGCCGATTTGGCCGCCTGGATTGTGTTGTCAACAAC
GCTGGCCACCACCCACCCCCACAGAGGCCTGAGGAGACCTCTGCCCAGGGATTCCGCCAGCT
GCTGGAGCTGAACCTACTGGGGACGTACACCTTGACCAAGCTCGCCCTCCCCTACCTGCGGA
AGAGTCAAGGGAATGTCATCAACATCTCCAGCCTGGTGGGGGCAATCGGCCAGGCCCAGGCA
GTTCCCTATGTGGCCACCAAGGGGGCAGTAACAGCCATGACCAAAGCTTTGGCCCTGGATGA
AAGTCCATATGGTGTCCGAGTCAACTGTATCTCCCCAGGAAACATCTGGACCCCGCTGTGGG
AGGAGCTGGCAGCCTTAATGCCAGACCCTAGGGCCACAATCCGAGAGGGCATGCTGGCCCAG
CCACTGGGCCGCATGGGCCAGCCCGCTGAGGTCGGGGCTGCGGCAGTGTTCCTGGCCTCCGA
AGCCAACTTCTGCACGGGCATTGAACTGCTCGTGACGGGGGGTGCAGAGCTGGGGTACGGGT
GCAAGGCCAGTCGGAGCACCCCCGTGGACGCCCCCGATATCCCTTCCTGATTTCTCTCATTT
CTACTTGGGGCCCCCTTCCTAGGACTCTCCCACCCCAAACTCCAACCTGTATCAGATGCAGC
CCCCAAGCCCTTAGACTCTAAGCCCAGTTAGCAAGGTGCCGGGTCACCCTGCAGGTTCCCAT
AAAAACGATTTGCAGCC

FIGURE 196

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56045
<subunit 1 of 1, 270 aa, 1 stop
<MW: 28317, pI: 6.00, NX(S/T): 1
MATGTRYAGKVVVVTGGGRGIGAGIVRAFVNSGARVVICDKDESGGRALEQELPGAVFILCD
VTQEDDVKTLVSETIRRFGRLDCVVNNAGHHPPPQRPEETSAQGFRQLLELNLLGTYTLTKL
ALPYLRKSQGNVINISSLVGAIGQAQAVPYVATKGAVTAMTKALALDESPYGVRVNCISPGN
IWTPLWEELAALMPDPRATIREGMLAQPLGRMGQPAEVGAAAVFLASEANFCTGIELLVTGG
AELGYGCKASRSTPVDAPDIPS

Important features:

N-glycosylation site.

amino acids 138-141

Short-chain alcohol dehydrogenase family protein

amino acids 10-22, 81-91, 134-171 and 176-185

FIGURE 197

AGGCGGGCAGCAGCTGCAGGCTGACCTTGCAGCTTGGCGGAATGGACTGGCCTCACAACCTG
CTGTTTCTTCTTACCATTTCCATCTTCCTGGGGCTGGGCCAGCCCAGGAGCCCCAAAAGCAA
GAGGAAGGGGCAAGGGCGGCCTGGGCCCCTGGCCCCTGGCCCTCACCAGGTGCCACTGGACC
TGGTGTCACGGATGAAACCGTATGCCCGCATGGAGGAGTATGAGAGGAACATCGAGGAGATG
GTGGCCCAGCTGAGGAACAGCTCAGAGCTGGCCCAGAGAAAGTGTGAGGTCAACTTGCAGCT
GTGGATGTCCAACAAGAGGAGCCTGTCTCCCTGGGGCTACAGCATCAACCACGACCCCAGCC
GTATCCCCGTGGACCTGCCGGAGGCACGGTGCCTGTGTCTGGGCTGTGTGAACCCCTTCACC
ATGCAGGAGGACCGCAGCATGGTGAGCGTGCCGGTGTTCAGCCAGGTTCCTGTGCGCCGCCG
CCTCTGCCCGCCACCGCCCCGCACAGGGCCTTGCCGCCAGCGCGCAGTCATGGAGACCATCG
CTGTGGGCTGCACCTGCATCTTCTGAATCACCTGGCCCAGAAGCCAGGCCAGCAGCCCGAGA
CCATCCTCCTTGCACCTTTGTGCCAAGAAAGGCCTATGAAAAGTAAACACTGACTTTTGAAA
GCAAG

FIGURE 198

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59294
<subunit 1 of 1, 180 aa, 1 stop
<MW: 20437, pI: 9.58, NX(S/T): 1
MDWPHNLLFLLTISIFLGLGQPRSPKSKRKGQGRPGPLAPGPHQVPLDLVSRMKPYARMEEY
ERNIEEMVAQLRNSSELAQRKCEVNLQLWMSNKRSLSPWGYSINHDPSRIPVDLPEARCLCL
GCVNPFTMQEDRSMVSVPVFSQVPVRRRLCPPPPRTGPCRQRAVMETIAVGCTCIF

Important features:

Signal peptide:
amino acids 1-20

N-glycosylation site.
amino acids 75-78

Homologous region to IL-17
amino acids 96-180.

FIGURE 199

GCGCCGCCAGGCGTAGGCGGGGTGGCCCTTGCGTCTCCCGCTTCCTTGAAAAACCCGGCGGG
CGAGCGAGGCTGCGGGCCGGCCGCTGCCCTTCCCCACACTCCCCGCCGAGAAGCCTCGCTCG
GCGCCCAAC**ATG**GCGGGTGGGCGCTGCGGCCCGCAGCTAACGGCGCTCCTGGCCGCCTGGAT
CGCGGCTGTGGCGGCGACGGCAGGCCCCGAGGAGGCCGCGCTGCCGCCGGAGCAGAGCCGGG
TCCAGCCCATGACCGCCTCCAACTGGACGCTGGTGATGGAGGGCGAGTGGATGCTGAAATTT
TACGCCCCATGGTGTCCATCCTGCCAGCAGACTGATTCAGAATGGGAGGCTTTTGCAAAGAA
TGGTGAAATACTTCAGATCAGTGTGGGGAAGGTAGATGTCATTCAAGAACCAGGTTTGAGTG
GCCGCTTCTTTGTCACCACTCTCCCAGCATTTTTTCATGCAAAGGATGGGATATTCCGCCGT
TATCGTGGCCCAGGAATCTTCGAAGACCTGCAGAATTATATCTTAGAGAAGAAATGGCAATC
AGTCGAGCCTCTGACTGGCTGGAAATCCCCAGCTTCTCTAACGATGTCTGGAATGGCTGGTC
TTTTTAGCATCTCTGGCAAGATATGGCATCTTCACAACTATTTCACAGTGACTCTTGGAATT
CCTGCTTGGTGTTCTTATGTGTTTTTCGTCATAGCCACCTTGGTTTTTGGCCTTTTTATGGG
TCTGGTCTTGGTGGTAATATCAGAATGTTTCTATGTGCCACTTCCAAGGCATTTATCTGAGC
GTTCTGAGCAGAATCGGAGATCAGAGGAGGCTCATAGAGCTGAACAGTTGCAGGATGCGGAG
GAGGAAAAAGATGATTCAAATGAAGAAGAAAACAAAGACAGCCTTGTAGATGATGAAGAAGA
GAAAGAAGATCTTGGCGATGAGGATGAAGCAGAGGAAGAAGAGGAGGAGGACAACTTGGCTG
CTGGTGTGGATGAGGAGAGAAGTGAGGCCAATGATCAGGGGCCCCCAGGAGAGGACGGTGTG
ACCCGGGAGGAAGTAGAGCCTGAGGAGGCTGAAGAAGGCATCTCTGAGCAACCCTGCCCAGC
TGACACAGAGGTGGTGGAAGACTCCTTGAGGCAGCGTAAAAGTCAGCATGCTGACAAGGGAC
TG**TAG**ATTTAATGATGCGTTTTCAAGAATACACACCAAAACAATATGTCAGCTTCCCTTTGG
CCTGCAGTTTGTACCAAATCCTTAATTTTTCCTGAATGAGCAAGCTTCTCTTAAAAGATGCT
CTCTAGTCATTTGGTCTCATGGCAGTAAGCCTCATGTATACTAAGGAGAGTCTTCCAGGTGT
GACAATCAGGATATAGAAAAACAAACGTAGTGTTGGGATCTGTTTGGAGACTGGGATGGGAA
CAAGTTCATTTACTTAGGGGTCAGAGAGTCTCGACCAGAGGAGGCCATTCCCAGTCCTAATC
AGCACCTTCCAGAGACAAGGCTGCAGGCCCTGTGAAATGAAAGCCAAGCAGGAGCCTTGGCT
CCTGAGCATCCCCAAAGTGTAACGTAGAAGCCTTGCATCCTTTTCTTGTGTAAAGTATTTAT
TTTTGTCAAATTGCAGGAAACATCAGGCACCACAGTGCATGAAAAATCTTTCACAGCTAGAA
ATTGAAAGGGCCTTGGGTATAGAGAGCAGCTCAGAAGTCATCCCAGCCCTCTGAATCTCCTG
TGCTATGTTTTATTTCTTACCTTTAATTTTTCCAGCATTTCCACCATGGGCATTCAGGCTCT
CCACACTCTTCACTATTATCTCTTGGTCAGAGGACTCCAATAACAGCCAGGTTTACATGAAC
TGTGTTTGTTCATTCTGACCTAAGGGGTTTAGATAATCAGTAACCATAACCCCTGAAGCTGT
GACTGCCAAACATCTCAAATGAAATGTTGTGGCCATCAGAGACTCAAAAGGAAGTAAGGATT
TTACAAGACAGATTAAAAAAAAAATTGTTTTGTCCAAAATATAGTTGTTGTTGATTTTTTTTT
AAGTTTTCTAAGCAATATTTTTCAAGCCAGAAGTCCTCTAAGTCTTGCCAGTACAAGGTAGT
CTTGTGAAGAAAAGTTGAATACTGTTTTGTTTTCATCTCAAGGGGTTCCCTGGGTCTTGAAC
TACTTTAATAATAACTAAAAAACCACTTCTGATTTTCCTTCAGTGATGTGCTTTTGGTGAAA
GAATTAATGAACTCCAGTACCTGAAAGTGAAAGATTTGATTTTGTTTCCATCTTCTGTAATC
TTCCAAAGAATTATATCTTTGTAAATCTCTCAATACTCAATCTACTGTAAGTACCCAGGGAG
GCTAATTTCTTT

FIGURE 200

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56433
<subunit 1 of 1, 349 aa, 1 stop
<MW: 38952, pI: 4.34, NX(S/T): 1
MAGGRCGPQLTALLAAWIAAVAATAGPEEAALPPEQSRVQPMTASNWTLVMEGEWMLKFYAP
WCPSCQQTDSEWEAFAKNGEILQISVGKVDVIQEPGLSGRFFVTTLPAFFHAKDGIFRRYRG
PGIFEDLQNYILEKKWQSVEPLTGWKSPASLTMSGMAGLFSISGKIWHLHNYFTVTLGIPAW
CSYVFFVIATLVFGLFMGLVLVVISECFYVPLPRHLSERSEQNRRSEEAHRAEQLQDAEEEK
DDSNEEENKDSLVDDEEEKEDLGDEDEAEEEEEEDNLAAGVDEERSEANDQGPPGEDGVTRE
EVEPEEAEEGISEQPCPADTEVVEDSLRQRKSQHADKGL

Important features:

Signal peptide:
amino acids 1-22

Transmembrane domain:
amino acids 191-211

N-glycosylation site.
amino acids 46-49

Thioredoxin family proteins. (homologous region to disulfide isomerase)
amino acids 56-72

Flavodoxin proteins
amino acids 173-187

FIGURE 201

ATCTGGTTGAACTACTTAAGCTTAATTTGTTAAACTCCGGTAAGTACCTAGCCCACATGATT
TGACTCAGAGATTCTCTTTTGTCCACAGACAGTCATCTCAGGGGCAGAAAGAAAAGAGCTCC
CAAATGCTATATCTATTCAGGGGCTCTCAAGAACAATGGAATATCATCCTGATTTAGAAAAT
TTGGATGAAGATGGATATACTCAATTACACTTCGACTCTCAAAGCAATACCAGGATAGCTGT
TGTTTCAGAGAAAGGATCGTGTGCTGCATCTCCTCCTTGGCGCCTCATTGCTGTAATTTTGG
GAATCCTATGCTTGGTAATACTGGTGATAGCTGTGGTCCTGGGTACCATGGGGGTTCTTTCC
AGCCCTTGTCCTCCTAATTGGATTATATATGAGAAGAGCTGTTATCTATTCAGCATGTCACT
AAATTCCTGGGATGGAAGTAAAAGACAATGCTGGCAACTGGGCTCTAATCTCCTAAAGATAG
ACAGCTCAAATGAATTGGGATTTATAGTAAAACAAGTGTCTTCCCAACCTGATAATTCATTT
TGGATAGGCCTTTCTCGGCCCCAGACTGAGGTACCATGGCTCTGGGAGGATGGATCAACATT
CTCTTCTAACTTATTTCAGATCAGAACCACAGCTACCCAAGAAAACCCATCTCCAAATTGTG
TATGGATTCACGTGTCAGTCATTTATGACCAACTGTGTAGTGTGCCCTCATATAGTATTTGT
GAGAAGAAGTTTTCAATGTAAGAGGAAGGGTGGAGAAGGAGAGAGAAATATGTGAGGTAGTA
AGGAGGACAGAAAACAGAACAGAAAAGAGTAACAGCTGAGGTCAAGATAAATGCAGAAAATG
TTTAGAGAGCTTGGCCAACTGTAATCTTAACCAAGAAATTGAAGGGAGAGGCTGTGATTTCT
GTATTTGTCGACCTACAGGTAGGCTAGTATTATTTTTCTAGTTAGTAGATCCCTAGACATGG
AATCAGGGCAGCCAAGCTTGAGTTTTTATTTTTTATTTATTTATTTTTTTGAGATAGGGTCT
CACTTTGTTACCCAGGCTGGAGTGCAGTGGCACAATCTCGACTCACTGCAGCTATCTCTCGC
CTCAGCCCCTCAAGTAGCTGGGACTACAGGTGCATGCCACCATGCCAGGCTAATTTTTGGTG
TTTTTTGTAGAGACTGGGTTTTGCCATGTTGACCAAGCTGGTCTCTAACTCCTGGGCTTAAG
TGATCTGCCCGCCTTGGCCTCCCAAAGTGCTGGGATTACAGATGTGAGCCACCACACCTGGC
CCCAAGCTTGAATTTTCATTCTGCCATTGACTTGGCATTTACCTTGGGTAAGCCATAAGCGA
ATCTTAATTTCTGGCTCTATCAGAGTTGTTTCATGCTCAACAATGCCATTGAAGTGCACGGT
GTGTTGCCACGATTTGACCCTCAACTTCTAGCAGTATATCAGTTATGAACTGAGGGTGAAAT
ATATTTCTGAATAGCTAAATGAAGAAATGGGAAAAAATCTTCACCACAGTCAGAGCAATTTT
ATTATTTTCATCAGTATGATCATAATTATGATTATCATCTTAGTAAAAAGCAGGAACTCCTA
CTTTTTCTTTATCAATTAAATAGCTCAGAGAGTACATCTGCCATATCTCTAATAGAATCTTT
TTTTTTTTTTTTTTTTTTTGAGACAGAGTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAACGG
CACGATCTCGGCTCACCGCAACCTCCGCCCCCTGGGTTCAAGCAATTCTCCTGCCTCAGCCT
CCCAAGTAGCTGGGATTACAGTCAGGCACCACCACACCCGGCTAATTTTGTATTTTTTTAGT
AGAGACAGGGTTTCTCCATGTCGGTCAGGGTAGTCCCGAACTCCTGACCTCAAGTGATCTGC
CTGCCTCGGCCTCCCAAGTGCTGGGATTACAGGCGTGAGCCACTGCACCCAGCCTAGAATCT
TGTATAATATGTAATTGTAGGGAAACTGCTCTCATAGGAAAGTTTTCTGCTTTTTAAATACA
AAAATACATAAAAATACATAAAATCTGATGATGAATATAAAAAAGTAACCAACCTCATTGGA
ACAAGTATTAACATTTTGGAATATGTTTTATTAGTTTTGTGATGTACTGTTTTACAATTTTT
ACCATTTTTTTCAGTAATTACTGTAAAATGGTATTATTGGAATGAAACTATATTTCCTCATG
TGCTGATTTGTCTTATTTTTTTCATACTTTCCCACTGGTGCTATTTTTATTTCCAATGGATA
TTTCTGTATTACTAGGGAGGCATTTACAGTCCTCTAATGTTGATTAATATGTGAAAAGAAAT
TGTACCAATTTTACTAAATTATGCAGTTTAAAATGGATGATTTTATGTTATGTGGATTTCAT
TTCAATAAAAAAAAACTCTTATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 202

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53912
<subunit 1 of 1, 201 aa, 1 stop
<MW: 22563, pI: 4.87, NX(S/T): 1
MEYHPDLENLDEDGYTQLHFDSQSNTRIAVVSEKGSCAASPPWRLIAVILGILCLVILVIAV
VLGTMGVLSSPCPPNWIIYEKSCYLFSMSLNSWDGSKRQCWQLGSNLLKIDSSNELGFIVKQ
VSSQPDNSFWIGLSRPQTEVPWLWEDGSTFSSNLFQIRTTATQENPSPNCVWIHVSVIYDQL
CSVPSYSICEKKFSM

Important features:

Type II transmembrane domain:
amino acids 45-65

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 197-200

N-myristoylation sites.
amino acids 35-40 and 151-156

Homologous region to LDL receptor
amino acids 34-67 and 70-200.

FIGURE 203A

GGAAGGGGAGGAGCAGGCCACACAGGCACAGGCCGGTGAGGGACCTGCCCAGACCTGGAGGG
TCTCGCTCTGTCACACAGGCTGGAGTGCAGTGGTGTGATCTTGGCTCATCGTAACCTCCACC
TCCCGGGTTCAAGTGATTCTCATGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGTGGTGAC
TTCCAAGAGTGACTCCGTCGGAGGAAAATGACTCCCCAGTCGCTGCTGCAGACGACACTGTT
CCTGCTGAGTCTGCTCTTCCTGGTCCAAGGTGCCCACGGCAGGGGCCACAGGGAAGACTTTC
GCTTCTGCAGCCAGCGGAACCAGACACACAGGAGCAGCCTCCACTACAAACCCACACCAGAC
CTGCGCATCTCCATCGAGAACTCCGAAGAGGCCCTCACAGTCCATGCCCCTTTCCCTGCAGC
CCACCCTGCTTCCCGATCCTTCCCTGACCCCAGGGGCCTCTACCACTTCTGCCTCTACTGGA
ACCGACATGCTGGGAGATTACATCTTCTCTATGGCAAGCGTGACTTCTTGCTGAGTGACAAA
GCCTCTAGCCTCCTCTGCTTCCAGCACCAGGAGGAGAGCCTGGCTCAGGGCCCCCCGCTGTT
AGCCACTTCTGTCACCTCCTGGTGGAGCCCTCAGAACATCAGCCTGCCCAGTGCCGCCAGCT
TCACCTTCTCCTTCCACAGTCCTCCCCACACGGCCGCTCACAATGCCTCGGTGGACATGTGC
GAGCTCAAAAGGGACCTCCAGCTGCTCAGCCAGTTCCTGAAGCATCCCCAGAAGGCCTCAAG
GAGGCCCTCGGCTGCCCCCGCCAGCCAGCAGTTGCAGAGCCTGGAGTCGAAACTGACCTCTG
TGAGATTCATGGGGGACATGGTGTCCTTCGAGGAGGACCGGATCAACGCCACGGTGTGGAAG
CTCCAGCCCACAGCCGGCCTCCAGGACCTGCACATCCACTCCCGGCAGGAGGAGGAGCAGAG
CGAGATCATGGAGTACTCGGTGCTGCTGCCTCGAACACTCTTCCAGAGGACGAAAGGCCGGA
GCGGGGAGGCTGAGAAGAGACTCCTCCTGGTGGACTTCAGCAGCCAAGCCCTGTTCCAGGAC
AAGAATTCCAGCCAAGTCCTGGGTGAGAAGGTCTTGGGGATTGTGGTACAGAACACCAAAGT
AGCCAACCTCACGGAGCCCGTGGTGCTCACTTTCCAGCACCAGCTACAGCCGAAGAATGTGA
CTCTGCAATGTGTGTTCTGGGTTGAAGACCCCACATTGAGCAGCCCGGGGCATTGGAGCAGT
GCTGGGTGTGAGACCGTCAGGAGAGAAACCCAAACATCCTGCTTCTGCAACCACTTGACCTA
CTTTGCAGTGCTGATGGTCTCCTCGGTGGAGGTGGACGCCGTGCACAAGCACTACCTGAGCC
TCCTCTCCTACGTGGGCTGTGTCGTCTCTGCCCTGGCCTGCCTTGTCACCATTGCCGCCTAC
CTCTGCTCCAGGGTGCCCCTGCCGTGCAGGAGGAAACCTCGGGACTACACCATCAAGGTGCA
CATGAACCTGCTGCTGGCCGTCTTCCTGCTGGACACGAGCTTCCTGCTCAGCGAGCCGGTGG
CCCTGACAGGCTCTGAGGCTGGCTGCCGAGCCAGTGCCATCTTCCTGCACTTCTCCCTGCTC
ACCTGCCTTTCCTGGATGGGCCTCGAGGGGTACAACCTCTACCGACTCGTGGTGGAGGTCTT
TGGCACCTATGTCCCTGGCTACCTACTCAAGCTGAGCGCCATGGGCTGGGGCTTCCCCATCT
TTCTGGTGACGCTGGTGGCCCTGGTGGATGTGGACAACTATGGCCCCATCATCTTGGCTGTG
CATAGGACTCCAGAGGGCGTCATCTACCCTTCCATGTGCTGGATCCGGGACTCCCTGGTCAG
CTACATCACCAACCTGGGCCTCTTCAGCCTGGTGTTTCTGTTCAACATGGCCATGCTAGCCA
CCATGGTGGTGCAGATCCTGCGGCTGCGCCCCCACACCCAAAAGTGGTCACATGTGCTGACA
CTGCTGGGCCTCAGCCTGGTCCTTGGCCTGCCCTGGGCCTTGATCTTCTTCTCCTTTGCTTC
TGGCACCTTCCAGCTTGTCGTCCTCTACCTTTTCAGCATCATCACCTCCTTCCAAGGCTTCC
TCATCTTCATCTGGTACTGGTCCATGCGGCTGCAGGCCCGGGGTGGCCCCTCCCCTCTGAAG
AGCAACTCAGACAGCGCCAGGCTCCCCATCAGCTCGGGCAGCACCTCGTCCAGCCGCATCTA
GGCCTCCAGCCCACCTGCCCATGTGATGAAGCAGAGATGCGGCCTCGTCGCACACTGCCTGT
GGCCCCCGAGCCAGGCCCAGCCCCAGGCCAGTCAGCCGCAGACTTTGGAAAGCCCAACGACC
ATGGAGAGATGGGCCGTTGCCATGGTGGACGGACTCCCGGGCTGGGCTTTTGAATTGGCCTT
GGGGACTACTCGGCTCTCACTCAGCTCCCACGGGACTCAGAAGTGCGCCGCCATGCTGCCTA
GGGTACTGTCCCCACATCTGTCCCAACCCAGCTGGAGGCCTGGTCTCTCCTTACAACCCCTG
GGCCCAGCCCTCATTGCTGGGGGCCAGGCCTTGGATCTTGAGGGTCTGGCACATCCTTAATC
CTGTGCCCCTGCCTGGGACAGAAATGTGGCTCCAGTTGCTCTGTCTCTCGTGGTCACCCTGA
GGGCACTCTGCATCCTCTGTCATTTTAACCTCAGGTGGCACCCAGGGCGAATGGGGCCCAGG
GCAGACCTTCAGGGCCAGAGCCCTGGCGGAGGAGAGGCCCTTTGCCAGGAGCACAGCAGCAG
CTCGCCTACCTCTGAGCCCAGGCCCCCTCCCTCCCTCAGCCCCCCAGTCCTCCCTCCATCTT
CCCTGGGGTTCTCCTCCTCTCCCAGGGCCTCCTTGCTCCTTCGTTCACAGCTGGGGGTCCCC
GATTCCAATGCTGTTTTTTGGGGAGTGGTTTCCAGGAGCTGCCTGGTGTCTGCTGTAAATGT
TTGTCTACTGCACAAGCCTCGGCCTGCCCCTGAGCCAGGCTCGGTACCGATGCGTGGGCTGG
GCTAGGTCCCTCTGTCCATCTGGGCCTTTGTATGAGCTGCATTGCCCTTGCTCACCCTGACC
AAGCACACGCCTCAGAGGGGCCCTCAGCCTCTCCTGAAGCCCTCTTGTGGCAAGAACTGTGG
ACCATGCCAGTCCCGTCTGGTTTCCATCCCACCACTCCAAGGACTGAGACTGACCTCCTCTG
GTGACACTGGCCTAGAGCCTGACACTCTCCTAAGAGGTTCTCTCCAAGCCCCCAAATAGCTC

FIGURE 203B

CAGGCGCCCTCGGCCGCCCATCATGGTTAATTCTGTCCAACAAACACACACGGGTAGATTGC
TGGCCTGTTGTAGGTGGTAGGGACACAGATGACCGACCTGGTCACTCCTCCTGCCAACATTC
AGTCTGGTATGTGAGGCGTGCGTGAAGCAAGAACTCCTGGAGCTACAGGGACAGGGAGCCAT
CATTCCTGCCTGGGAATCCTGGAAGACTTCCTGCAGGAGTCAGCGTTCAATCTTGACCTTGA
AGATGGGAAGGATGTTCTTTTTACGTACCAATTCTTTTGTCTTTTGATATTAAAAAGAAGTA
CATGTTCATTGTAGAGAATTTGGAAACTGTAGAAGAGAATCAAGAAGAAAAATAAAAATCAG
CTGTTGTAATCGCCTAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 204

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50921
<subunit 1 of 1, 693 aa, 1 stop
<MW: 77738, pI: 8.87, NX(S/T): 7
MTPQSLLQTTLFLLSLLFLVQGAHGRGHREDFRFCSQRNQTHRSSLHYKPTPDLRISIENSE
EALTVHAPFPAAHPASRSFPDPRGLYHFCLYWNRHAGRLHLLYGKRDFLLSDKASSLLCFQH
QEESLAQGPPLLATSVTSWWSPQNISLPSAASFTFSFHSPPHTAAHNASVDMCELKRDLQLL
SQFLKHPQKASRRPSAAPASQQLQSLESKLTSVRFMGDMVSFEEDRINATVWKLQPTAGLQD
LHIHSRQEEEQSEIMEYSVLLPRTLFQRTKGRSGEAEKRLLLVDFSSQALFQDKNSSQVLGE
KVLGIVVQNTKVANLTEPVVLTFQHQLQPKNVTLQCVFWVEDPTLSSPGHWSSAGCETVRRE
TQTSCFCNHLTYFAVLMVSSVEVDAVHKHYLSLLSYVGCVVSALACLVTIAAYLCSRVPLPC
RRKPRDYTIKVHMNLLLAVFLLDTSFLLSEPVALTGSEAGCRASAIFLHFSLLTCLSWMGLE
GYNLYRLVVEVFGTYVPGYLLKLSAMGWGFPIFLVTLVALVDVDNYGPIILAVHRTPEGVIY
PSMCWIRDSLVSYITNLGLFSLVFLFNMAMLATMVVQILRLRPHTQKWSHVLTLLGLSLVLG
LPWALIFFSFASGTFQLVVLYLFSIITSFQGFLIFIWYWSMRLQARGGPSPLKSNSDSARLP
ISSGSTSSSRI

Important features:

Signal peptide:
amino acids 1-25

Putative transmembrane domains:
amino acids 382-398, 402-420, 445-468, 473-491, 519-537, 568-590 and 634-657

Microbodies C-terminal targeting signal.
amino acids 691-693

cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 198-201 and 370-373

N-glycosylation sites.
amino acids 39-42, 148-151, 171-174, 234-237, 303-306, 324-327 and 341-344

G-protein coupled receptors family 2 proteins
amino acids 475-504

FIGURE 205

TGCCTGGCCTGCCTTGTCAACAATGCCGCTTACTCTGCTTCCAGGTTGCCCTGCCTTGCAGA
GGAAANCNTCGGGACTACACCNTCAAGTGCACATGAACCTGCTGCTGGCCGTCTTCCTGCTG
GACACGAGCTTCCTGCTCAGCGNAGCCGGTGGCCCTGACAGGCTCTGAAGGCTGGCTGCCGA
GCCAGTGCCATCTTCCTGCACTTCTCCTGCTCACCTGCCTTTCCTGGATGGGCCTCGAGGGG
TACAACCTCTACCGACTCGTGGTGGAGGTCTTTGGCACCTATGTCCCTGGCTACCTACTCAA
GCTGAGCGCCATGGGCTGGGGCTTCCCCATCTTTCTGGTGACGCTGGTGGCCCTGGTGGATG
TGGACAACTATGGCCCCATCATCTTGGCTGTGCATAGGACTCCAGAGGGCGTCATCTACCCT
TCCATGTGCTGGATCCGGGACTCCCTGGTCAGCTACATCACCAACCTGGGCCTCTTCAGCCT
GGTGTTTCTGTTCAACATGG

FIGURE 206

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCTGGTTCAGGTCCA
GGTTTTGCTTTGATCCTTTTCAAAAACTGGAGACACAGAAGAGGGCTCTAGGAAAAAGTTTT
GGATGGGATTATGTGGAAACTACCCTGCGATTCTCTGCTGCCAGAGCAGGCTCGGCGCTTCC
ACCCCAGTGCAGCCTTCCCCTGGCGGTGGTGAAAGAGACTCGGGAGTCGCTGCTTCCAAAGT
GCCCGCCGTGAGTGAGCTCTCACCCCAGTCAGCCAAATGAGCCTCTTCGGGCTTCTCCTGCT
GACATCTGCCCTGGCCGGCCAGAGACAGGGGACTCAGGCGGAATCCAACCTGAGTAGTAAAT
TCCAGTTTTCCAGCAACAAGGAACAGAACGGAGTACAAGATCCTCAGCATGAGAGAATTATT
ACTGTGTCTACTAATGGAAGTATTCACAGCCCAAGGTTTCCTCATACTTATCCAAGAAATAC
GGTCTTGGTATGGAGATTAGTAGCAGTAGAGGAAAATGTATGGATACAACTTACGTTTGATG
AAAGATTTGGGCTTGAAGACCCAGAAGATGACATATGCAAGTATGATTTTGTAGAAGTTGAG
GAACCCAGTGATGGAACTATATTAGGGCGCTGGTGTGGTTCTGGTACTGTACCAGGAAAACA
GATTTCTAAAGGAAATCAAATTAGGATAAGATTTGTATCTGATGAATATTTTCCTTCTGAAC
CAGGGTTCTGCATCCACTACAACATTGTCATGCCACAATTCACAGAAGCTGTGAGTCCTTCA
GTGCTACCCCCTTCAGCTTTGCCACTGGACCTGCTTAATAATGCTATAACTGCCTTTAGTAC
CTTGGAAGACCTTATTCGATATCTTGAACCAGAGAGATGGCAGTTGGACTTAGAAGATCTAT
ATAGGCCAACTTGGCAACTTCTTGGCAAGGCTTTTGTTTTTGGAAGAAAATCCAGAGTGGTG
GATCTGAACCTTCTAACAGAGGAGGTAAGATTATACAGCTGCACACCTCGTAACTTCTCAGT
GTCCATAAGGGAAGAACTAAAGAGAACCGATACCATTTTCTGGCCAGGTTGTCTCCTGGTTA
AACGCTGTGGTGGGAACTGTGCCTGTTGTCTCCACAATTGCAATGAATGTCAATGTGTCCCA
AGCAAAGTTACTAAAAAATACCACGAGGTCCTTCAGTTGAGACCAAAGACCGGTGTCAGGGG
ATTGCACAAATCACTCACCGACGTGGCCCTGGAGCACCATGAGGAGTGTGACTGTGTGTGCA
GAGGGAGCACAGGAGGATAGCCGCATCACCACCAGCAGCTCTTGCCCAGAGCTGTGCAGTGC
AGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCATCCTTAATCTCAGTTGTTTGCTT
CAAGGACCTTTCATCTTCAGGATTTACAGTGCATTCTGAAAGAGGAGACATCAAACAGAATT
AGGAGTTGTGCAACAGCTCTTTTGAGAGGAGGCCTAAAGGACAGGAGAAAAGGTCTTCAATC
GTGGAAAGAAAATTAAATGTTGTATTAAATAGATCACCAGCTAGTTTCAGAGTTACCATGTA
CGTATTCCACTAGCTGGGTTCTGTATTTCAGTTCTTTCGATACGGCTTAGGGTAATGTCAGT
ACAGGAAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGCTTAACTCTAAAGCTCC
ATGTCCTGGGCCTAAAATCGTATAAAATCTGGATTTTTTTTTTTTTTTGCTCATATTCAC
ATATGTAAACCAGAACATTCTATGTACTACAAACCTGGTTTTTAAAAAGGAACTATGTTGCT
ATGAATTAAACTTGTGTCATGCTGATAGGACAGACTGGATTTTTCATATTTCTTATTAAAAT
TTCTGCCATTTAGAAGAAGAGAACTACATTCATGGTTTGGAAGAGATAAACCTGAAAAGAAG
AGTGGCCTTATCTTCACTTTATCGATAAGTCAGTTTATTTGTTTCATTGTGTACATTTTTAT
ATTCTCCTTTTGACATTATAACTGTTGGCTTTTCTAATCTTGTTAAATATATCTATTTTTAC
CAAAGGTATTTAATATTCTTTTTTATGACAACTTAGATCAACTATTTTTAGCTTGGTAAATT
TTTCTAAACACAATTGTTATAGCCAGAGGAACAAAGATGATATAAAATATTGTTGCTCTGAC
AAAAATACATGTATTTCATTCTCGTATGGTGCTAGAGTTAGATTAATCTGCATTTTAAAAAA
CTGAATTGGAATAGAATTGGTAAGTTGCAAAGACTTTTTGAAAATAATTAAATTATCATATC
TTCCATTCCTGTTATTGGAGATGAAAATAAAAAGCAACTTATGAAAGTAGACATTCAGATCC
AGCCATTACTAACCTATTCCTTTTTTGGGGAAATCTGAGCCTAGCTCAGAAAAACATAAAGC
ACCTTGAAAAAGACTTGGCAGCTTCCTGATAAAGCGTGCTGTGCTGTGCAGTAGGAACACAT
CCTATTTATTGTGATGTTGTGGTTTTATTATCTTAAACTCTGTTCCATACACTTGTATAAAT
ACATGGATATTTTTATGTACAGAAGTATGTCTCTTAACCAGTTCACTTATTGTACTCTGGCA
ATTTAAAAGAAAATCAGTAAAATATTTTGCTTGTAAAATGCTTAATATNGTGCCTAGGTTAT
GTGGTGACTATTTGAATCAAAAATGTATTGAATCATCAAATAAAAGAATGTGGCTATTTTGG
GGAGAAAATTAAAAAAAAAAAAAAAAAAAAGGTTTAGGGATAACAGGGTAATGCGGCC

FIGURE 207

MSLFGLLLLTSALAGQRQGTQAESNLSSKFQFSSNKEQNGVQDPQHERIITVSTNGSIHSPR
FPHTYPRNTVLVWRLVAVEENVWIQLTFDERFGLEDPEDDICKYDFVEVEEPSDGTILGRWC
GSGTVPGKQISKGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQFTEAVSPSVLPPSALPLDLL
NNAITAFSTLEDLIRYLEPERWQLDLEDLYRPTWQLLGKAFVFGRKSRVVDLNLLTEEVRLY
SCTPRNFSVSIREELKRTDTIFWPGCLLVKRCGGNCACCLHNCNECQCVPSKVTKKYHEVLQ
LRPKTGVRGLHKSLTDVALEHHEECDCVCRGSTGG

FIGURE 208

CCCATCTCAAGCTGATCTTGGCACCTCTCATGCTCTGCTCTCTTCAACCAGACCTCTACATT
CCATTTTGGAAGAAGACTAAAA**ATG**GTGTTTCCAATGTGGACACTGAAGAGACAAATTCTTA
TCCTTTTTAACATAATCCTAATTTCCAAACTCCTTGGGGCTAGATGGTTTCCTAAAACTCTG
CCCTGTGATGTCACTCTGGATGTTCCAAAGAACCATGTGATCGTGGACTGCACAGACAAGCA
TTTGACAGAAATTCCTGGAGGTATTCCCACGAACACCACGAACCTCACCCTCACCATTAACC
ACATACCAGACATCTCCCCAGCGTCCTTTCACAGACTGGACCATCTGGTAGAGATCGATTTC
AGATGCAACTGTGTACCTATTCCACTGGGGTCAAAAAACAACATGTGCATCAAGAGGCTGCA
GATTAAACCCAGAAGCTTTAGTGGACTCACTTATTTAAAATCCCTTTACCTGGATGGAAACC
AGCTACTAGAGATACCGCAGGGCCTCCCGCCTAGCTTACAGCTTCTCAGCCTTGAGGCCAAC
AACATCTTTTCCATCAGAAAAGAGAATCTAACAGAACTGGCCAACATAGAAATACTCTACCT
GGGCCAAAACTGTTATTATCGAAATCCTTGTTATGTTTCATATTCAATAGAGAAAGATGCCT
TCCTAAACTTGACAAAGTTAAAAGTGCTCTCCCTGAAAGATAACAATGTCACAGCCGTCCCT
ACTGTTTTGCCATCTACTTTAACAGAACTATATCTCTACAACAACATGATTGCAAAAATCCA
AGAAGATGATTTTAATAACCTCAACCAATTACAAATTCTTGACCTAAGTGGAAATTGCCCTC
GTTGTTATAATGCCCCATTTCCTTGTGCGCCGTGTAAAAATAATTCTCCCCTACAGATCCCT
GTAAATGCTTTTGATGCGCTGACAGAATTAAAAGTTTTACGTCTACACAGTAACTCTCTTCA
GCATGTGCCCCCAAGATGGTTTAAGAACATCAACAAACTCCAGGAACTGGATCTGTCCCAAA
ACTTCTTGGCCAAAGAAATTGGGGATGCTAAATTTCTGCATTTTCTCCCCAGCCTCATCCAA
TTGGATCTGTCTTTCAATTTTGAACTTCAGGTCTATCGTGCATCTATGAATCTATCACAAGC
ATTTTCTTCACTGAAAAGCCTGAAAATTCTGCGGATCAGAGGATATGTCTTTAAAGAGTTGA
AAAGCTTTAACCTCTCGCCATTACATAATCTTCAAAATCTTGAAGTTCTTGATCTTGGCACT
AACTTTATAAAAATTGCTAACCTCAGCATGTTTAAACAATTTAAAAGACTGAAAGTCATAGA
TCTTTCAGTGAATAAAATATCACCTTCAGGAGATTCAAGTGAAGTTGGCTTCTGCTCAAATG
CCAGAACTTCTGTAGAAAGTTATGAACCCCAGGTCCTGGAACAATTACATTATTTCAGATAT
GATAAGTATGCAAGGAGTTGCAGATTCAAAAACAAAGAGGCTTCTTTCATGTCTGTTAATGA
AAGCTGCTACAAGTATGGGCAGACCTTGGATCTAAGTAAAAATAGTATATTTTTTGTCAAGT
CCTCTGATTTTCAGCATCTTTCTTTCCTCAAATGCCTGAATCTGTCAGGAAATCTCATTAGC
CAAACTCTTAATGGCAGTGAATTCCAACCTTTAGCAGAGCTGAGATATTTGGACTTCTCCAA
CAACCGGCTTGATTTACTCCATTCAACAGCATTTGAAGAGCTTCACAAACTGGAAGTTCTGG
ATATAAGCAGTAATAGCCATTATTTTCAATCAGAAGGAATTACTCATATGCTAAACTTTACC
AAGAACCTAAAGGTTCTGCAGAAACTGATGATGAACGACAATGACATCTCTTCCTCCACCAG
CAGGACCATGGAGAGTGAGTCTCTTAGAACTCTGGAATTCAGAGGAAATCACTTAGATGTTT
TATGGAGAGAAGGTGATAACAGATACTTACAATTATTCAAGAATCTGCTAAAATTAGAGGAA
TTAGACATCTCTAAAAATTCCCTAAGTTTCTTGCCTTCTGGAGTTTTTGATGGTATGCCTCC
AAATCTAAAGAATCTCTCTTTGGCCAAAAATGGGCTCAAATCTTTCAGTTGGAAGAAACTCC
AGTGTCTAAAGAACCTGGAAACTTTGGACCTCAGCCACAACCAACTGACCACTGTCCCTGAG
AGATTATCCAACTGTTCCAGAAGCCTCAAGAATCTGATTCTTAAGAATAATCAAATCAGGAG
TCTGACGAAGTATTTTCTACAAGATGCCTTCCAGTTGCGATATCTGGATCTCAGCTCAAATA
AAATCCAGATGATCCAAAAGACCAGCTTCCCAGAAAATGTCCTCAACAATCTGAAGATGTTG
CTTTTGCATCATAATCGGTTTCTGTGCACCTGTGATGCTGTGTGGTTTGTCTGGTGGGTTAA
CCATACGGAGGTGACTATTCCTTACCTGGCCACAGATGTGACTTGTGTGGGGCCAGGAGCAC
ACAAGGGCCAAAGTGTGATCTCCCTGGATCTGTACACCTGTGAGTTAGATCTGACTAACCTG
ATTCTGTTCTCACTTTCCATATCTGTATCTCTCTTTCTCATGGTGATGATGACAGCAAGTCA
CCTCTATTTCTGGGATGTGTGGTATATTTACCATTTCTGTAAGGCCAAGATAAAGGGGTATC
AGCGTCTAATATCACCAGACTGTTGCTATGATGCTTTTATTGTGTATGACACTAAAGACCCA
GCTGTGACCGAGTGGGTTTTGGCTGAGCTGGTGGCCAAACTGGAAGACCCAAGAGAGAAACA
TTTTAATTTATGTCTCGAGGAAAGGGACTGGTTACCAGGGCAGCCAGTTCTGGAAAACCTTT
CCCAGAGCATACAGCTTAGCAAAAAGACAGTGTTTGTGATGACAGACAAGTATGCAAAGACT
GAAAATTTTAAGATAGCATTTTACTTGTCCCATCAGAGGCTCATGGATGAAAAAGTTGATGT
GATTATCTTGATATTTCTTGAGAAGCCCTTTCAGAAGTCCAAGTTCCTCCAGCTCCGGAAAA
GGCTCTGTGGGAGTTCTGTCCTTGAGTGGCCAACAAACCCGCAAGCTCACCCATACTTCTGG
CAGTGTCTAAAGAACGCCCTGGCCACAGACAATCATGTGGCCTATAGTCAGGTGTTCAAGGA
AACGGTC**TAG**CCCTTCTTTGCAAAACACAACTGCCTAGTTTACCAAGGAGAGGCCTGGC

FIGURE 209

MVFPMWTLKRQILILFNIILISKLLGARWFPKTLPCDVTLDVPKNHVIVDCTDKHLTEIPGG
IPTNTTNLTLTINHIPDISPASFHRLDHLVEIDFRCNCVPIPLGSKNNMCIKRLQIKPRSFS
GLTYLKSLYLDGNQLLEIPQGLPPSLQLLSLEANNIFSIRKENLTELANIEILYLGQNCYYR
NPCYVSYSIEKDAFLNLTKLKVLSLKDNNVTAVPTVLPSTLTELYLYNNMIAKIQEDDFNNL
NQLQILDLSGNCPRCYNAPFPCAPCKNNSPLQIPVNAFDALTELKVLRLHSNSLQHVPPRWF
KNINKLQELDLSQNFLAKEIGDAKFLHFLPSLIQLDLSFNFELQVYRASMNLSQAFSSLKSL
KILRIRGYVFKELKSFNLSPLHNLQNLEVLDLGTNFIKIANLSMFKQFKRLKVIDLSVNKIS
PSGDSSEVGFCSNARTSVESYEPQVLEQLHYFRYDKYARSCRFKNKEASFMSVNESCYKYGQ
TLDLSKNSIFFVKSSDFQHLSFLKCLNLSGNLISQTLNGSEFQPLAELRYLDFSNNRLDLLH
STAFEELHKLEVLDISSNSHYFQSEGITHMLNFTKNLKVLQKLMMNDNDISSSTSRTMESES
LRTLEFRGNHLDVLWREGDNRYLQLFKNLLKLEELDISKNSLSFLPSGVFDGMPPNLKNLSL
AKNGLKSFSWKKLQCLKNLETLDLSHNQLTTVPERLSNCSRSLKNLILKNNQIRSLTKYFLQ
DAFQLRYLDLSSNKIQMIQKTSFPENVLNNLKMLLLHHNRFLCTCDAVWFVWWVNHTEVTIP
YLATDVTCVGPGAHKGQSVISLDLYTCELDLTNLILFSLSISVSLFLMVMMTASHLYFWDVW
YIYHFCKAKIKGYQRLISPDCCYDAFIVYDTKDPAVTEWVLAELVAKLEDPREKHFNLCLEE
RDWLPGQPVLENLSQSIQLSKKTVFVMTDKYAKTENFKIAFYLSHQRLMDEKVDVIILIFLE
KPFQKSKFLQLRKRLCGSSVLEWPTNPQAHPYFWQCLKNALATDNHVAYSQVFKETV

FIGURE 210A

GGGTACCATTCTGCGCTGCTGCAAGTTACGGAATGAAAAATTAGAACAACAGAAACATGGAA
AACATGTTCCTTCAGTCGTCAATGCTGACCTGCATTTTCCTGCTAATATCTGGTTCCTGTGA
GTTATGCGCCGAAGAAAATTTTTCTAGAAGCTATCCTTGTGATGAGAAAAAGCAAAATGACT
CAGTTATTGCAGAGTGCAGCAATCGTCGACTACAGGAAGTTCCCCAAACGGTGGGCAAATAT
GTGACAGAACTAGACCTGTCTGATAATTTCATCACACACATAACGAATGAATCATTTCAAGG
GCTGCAAAATCTCACTAAAATAAATCTAAACCACAACCCCAATGTACAGCACCAGAACGGAA
ATCCCGGTATACAATCAAATGGCTTGAATATCACAGACGGGGCATTCCTCAACCTAAAAAAC
CTAAGGGAGTTACTGCTTGAAGACAACCAGTTACCCCAAATACCCTCTGGTTTGCCAGAGTC
TTTGACAGAACTTAGTCTAATTCAAAACAATATATACAACATAACTAAAGAGGGCATTTCAA
GACTTATAAACTTGAAAAATCTCTATTTGGCCTGGAACTGCTATTTTAACAAAGTTTGCGAG
AAAACTAACATAGAAGATGGAGTATTTGAAACGCTGACAAATTTGGAGTTGCTATCACTATC
TTTCAATTCTCTTTCACACGTGCCACCCAAACTGCCAAGCTCCCTACGCAAACTTTTTCTGA
GCAACACCCAGATCAAATACATTAGTGAAGAAGATTTCAAGGGATTGATAAATTTAACATTA
CTAGATTTAAGCGGGAACTGTCCGAGGTGCTTCAATGCCCCATTTCCATGCGTGCCTTGTGA
TGGTGGTGCTTCAATTAATATAGATCGTTTTGCTTTTCAAAACTTGACCCAACTTCGATACC
TAAACCTCTCTAGCACTTCCCTCAGGAAGATTAATGCTGCCTGGTTTAAAAATATGCCTCAT
CTGAAGGTGCTGGATCTTGAATTCAACTATTTAGTGGGAGAAATAGTCTCTGGGGCATTTTT
AACGATGCTGCCCCGCTTAGAAATACTTGACTTGTCTTTTAACTATATAAAGGGGAGTTATC
CACAGCATATTAATATTTCCAGAAACTTCTCTAAACTTTTGTCTCTACGGGCATTGCATTTA
AGAGGTTATGTGTTCCAGGAACTCAGAGAAGATGATTTCCAGCCCCTGATGCAGCTTCCAAA
CTTATCGACTATCAACTTGGGTATTAATTTTATTAAGCAAATCGATTTCAAACTTTTCCAAA
ATTTCTCCAATCTGGAAATTATTTACTTGTCAGAAAACAGAATATCACCGTTGGTAAAAGAT
ACCCGGCAGAGTTATGCAAATAGTTCCTCTTTTCAACGTCATATCCGGAAACGACGCTCAAC
AGATTTTGAGTTTGACCCACATTCGAACTTTTATCATTTCACCCGTCCTTTAATAAAGCCAC
AATGTGCTGCTTATGGAAAAGCCTTAGATTTAAGCCTCAACAGTATTTTCTTCATTGGGCCA
AACCAATTTGAAAATCTTCCTGACATTGCCTGTTTAAATCTGTCTGCAAATAGCAATGCTCA
AGTGTTAAGTGGAACTGAATTTTCAGCCATTCCTCATGTCAAATATTTGGATTTGACAAACA
ATAGACTAGACTTTGATAATGCTAGTGCTCTTACTGAATTGTCCGACTTGGAAGTTCTAGAT
CTCAGCTATAATTCACACTATTTCAGAATAGCAGGCGTAACACATCATCTAGAATTTATTCA
AAATTTCACAAATCTAAAAGTTTTAAACTTGAGCCACAACAACATTTATACTTTAACAGATA
AGTATAACCTGGAAAGCAAGTCCCTGGTAGAATTAGTTTTCAGTGGCAATCGCCTTGACATT
TTGTGGAATGATGATGACAACAGGTATATCTCCATTTTCAAAGGTCTCAAGAATCTGACACG
TCTGGATTTATCCCTTAATAGGCTGAAGCACATCCCAAATGAAGCATTCCTTAATTTGCCAG
CGAGTCTCACTGAACTACATATAAATGATAATATGTTAAAGTTTTTTAACTGGACATTACTC
CAGCAGTTTCCTCGTCTCGAGTTGCTTGACTTACGTGGAAACAAACTACTCTTTTTAACTGA
TAGCCTATCTGACTTTACATCTTCCCTTCGGACACTGCTGCTGAGTCATAACAGGATTTCCC
ACCTACCCTCTGGCTTTCTTTCTGAAGTCAGTAGTCTGAAGCACCTCGATTTAAGTTCCAAT
CTGCTAAAAACAATCAACAAATCCGCACTTGAAACTAAGACCACCACCAAATTATCTATGTT
GGAACTACACGGAAACCCCTTTGAATGCACCTGTGACATTGGAGATTTCCGAAGATGGATGG
ATGAACATCTGAATGTCAAAATTCCCAGACTGGTAGATGTCATTTGTGCCAGTCCTGGGGAT
CAAAGAGGGAAGAGTATTGTGAGTCTGGAGCTAACAACTTGTGTTTCAGATGTCACTGCAGT
GATATTATTTTTCTTCACGTTCTTTATCACCACCATGGTTATGTTGGCTGCCCTGGCTCACC
ATTTGTTTTACTGGGATGTTTGGTTTATATATAATGTGTGTTTAGCTAAGGTAAAAGGCTAC
AGGTCTCTTTCCACATCCCAAACTTTCTATGATGCTTACATTTCTTATGACACCAAAGATGC
CTCTGTTACTGACTGGGTGATAAATGAGCTGCGCTACCACCTTGAAGAGAGCCGAGACAAAA
ACGTTCTCCTTTGTCTAGAGGAGAGGGATTGGGACCCGGGATTGGCCATCATCGACAACCTC
ATGCAGAGCATCAACCAAAGCAAGAAAACAGTATTTGTTTTAACCAAAAAATATGCAAAAAG
CTGGAACTTTAAAACAGCTTTTTACTTGGCTTTGCAGAGGCTAATGGATGAGAACATGGATG
TGATTATATTTATCCTGCTGGAGCCAGTGTTACAGCATTCTCAGTATTTGAGGCTACGGCAG
CGGATCTGTAAGAGCTCCATCCTCCAGTGGCCTGACAACCCGAAGGCAGAAGGCTTGTTTTG
GCAAACTCTGAGAAATGTGGTCTTGACTGAAAATGATTCACGGTATAACAATATGTATGTCG
ATTCCATTAAGCAATACTAACTGACGTTAAGTCATGATTTCGCGCCATAATAAAGATGCAAA
GGAATGACATTTCTGTATTAGTTATCTATTGCTATGTAACAAATTATCCCAAAACTTAGTGG
TTTAAAACAACACATTTGCTGGCCCACAGTTTTTGAGGGTCAGGAGTCCAGGCCCAGCATAA

FIGURE 210B

CTGGGTCCTCTGCTCAGGGTGTCTCAGAGGCTGCAATGTAGGTGTTCACCAGAGACATAGGC
ATCACTGGGGTCACACTCATGTGGTTGTTTTCTGGATTCAATTCCTCCTGGGCTATTGGCCA
AAGGCTATACTCATGTAAGCCATGCGAGCCTCTCCCACAAGGCAGCTTGCTTCATCAGAGCT
AGCAAAAAAGAGAGGTTGCTAGCAAGATGAAGTCACAATCTTTTGTAATCGAATCAAAAAAG
TGATATCTCATCACTTTGGCCATATTCTATTTGTTAGAAGTAAACCACAGGTCCCACCAGCT
CCATGGGAGTGACCACCTCAGTCCAGGGAAAACAGCTGAAGACCAAGATGGTGAGCTCTGAT
TGCTTCAGTTGGTCATCAACTATTTTCCCTTGACTGCTGTCCTGGGATGGCCTGCTATCTTG
ATGATAGATTGTGAATATCAGGAGGCAGGGATCACTGTGGACCATCTTAGCAGTTGACCTAA
CACATCTTCTTTTCAATATCTAAGAACTTTTGCCACTGTGACTAATGGTCCTAATATTAAGC
TGTTGTTTATATTTATCATATATCTATGGCTACATGGTTATATTATGCTGTGGTTGCGTTCG
GTTTTATTTACAGTTGCTTTTACAAATATTTGCTGTAACATTTGACTTCTAAGGTTTAGATG
CCATTTAAGAACTGAGATGGATAGCTTTTAAAGCATCTTTTACTTCTTACCATTTTTTAAAA
GTATGCAGCTAAATTCGAAGCTTTTGGTCTATATTGTTAATTGCCATTGCTGTAAATCTTAA
AATGAATGAATAAAAATGTTTCATTTTACAAAAAAAAAAAAAAA

FIGURE 211

MENMFLQSSMLTCIFLLISGSCELCAEENFSRSYPCDEKKQNDSVIAECSNRRLQEVPQTVG
KYVTELDLSDNFITHITNESFQGLQNLTKINLNHNPNVQHQNGNPGIQSNGLNITDGAFLNL
KNLRELLLEDNQLPQIPSGLPESLTELSLIQNNIYNITKEGISRLINLKNLYLAWNCYFNKV
CEKTNIEDGVFETLTNLELLSLSFNSLSHVPPKLPSSLRKLFLSNTQIKYISEEDFKGLINL
TLLDLSGNCPRCFNAPFPCVPCDGGASINIDRFAFQNLTQLRYLNLSSTSLRKINAAWFKNM
PHLKVLDLEFNYLVGEIVSGAFLTMLPRLEILDLSFNYIKGSYPQHINISRNFSKLLSLRAL
HLRGYVFQELREDDFQPLMQLPNLSTINLGINFIKQIDFKLFQNFSNLEIIYLSENRISPLV
KDTRQSYANSSSFQRHIRKRRSTDFEFDPHSNFYHFTRPLIKPQCAAYGKALDLSLNSIFFI
GPNQFENLPDIACLNLSANSNAQVLSGTEFSAIPHVKYLDLTNNRLDFDNASALTELSDLEV
LDLSYNSHYFRIAGVTHHLEFIQNFTNLKVLNLSHNNIYTLTDKYNLESKSLVELVFSGNRL
DILWNDDDNRYISIFKGLKNLTRLDLSLNRLKHIPNEAFLNLPASLTELHINDNMLKFFNWT
LLQQFPRLELLDLRGNKLLFLTDSLSDFTSSLRTLLLSHNRISHLPSGFLSEVSSLKHLDLS
SNLLKTINKSALETKTTTKLSMLELHGNPFECTCDIGDFRRWMDEHLNVKIPRLVDVICASP
GDQRGKSIVSLELTTCVSDVTAVILFFFTFFITTMVMLAALAHHLFYWDVWFIYNVCLAKVK
GYRSLSTSQTFYDAYISYDTKDASVTDWVINELRYHLEESRDKNVLLCLEERDWDPGLAIID
NLMQSINQSKKTVFVLTKKYAKSWNFKTAFYLALQRLMDENMDVIIFILLEPVLQHSQYLRL
RQRICKSSILQWPDNPKAEGLFWQTLRNVVLTENDSRYNNMYVDSIKQY

FIGURE 212

CCAGGTCCAACTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCT
CGACCTCGACCCACGCGTCCGCCAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGA
CAGGCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGC
AAGGGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGC
TCCAGCAGCATCAGCAGCCCCCAGGACCGGGGAGGCACAGGTGGCCCCCACCACCCGGAGGA
GCAGCTCCTGCCCCTGTCCGGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGAA
GGCCACCCCGCCTGGAGGCACAGGCCATGAGGGGCTCTCAGGAGGTGCTGCTGATGTGGCTT
CTGGTGTTGGCAGTGGGCGGCACAGAGCACGCCTACCGGCCCGGCCGTAGGGTGTGTGCTGT
CCGGGCTCACGGGGACCCTGTCTCCGAGTCGTTCGTGCAGCGTGTGTACCAGCCCTTCCTCA
CCACCTGCGACGGGCACCGGGCCTGCAGCACCTACCGAACCATCTATAGGACCGCCTACCGC
CGCAGCCCTGGGCTGGCCCCTGCCAGGCCTCGCTACGCGTGCTGCCCCGGCTGGAAGAGGAC
CAGCGGGCTTCCTGGGGCCTGTGGAGCAGCAATATGCCAGCCGCCATGCCGGAACGGAGGGA
GCTGTGTCCAGCCTGGCCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAGTCA
GATGTGGATGAATGCAGTGCTAGGAGGGGCGGCTGTCCCCAGCGCTGCATCAACACCGCCGG
CAGTTACTGGTGCCAGTGTTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTGTGC
CCAAGGGAGGGCCCCCCAGGGTGGCCCCCAACCCGACAGGAGTGGACAGTGCAATGAAGGAA
GAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGCTGGC
CCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGCATGGGCTCCCGGACCCCGGCAGCCTCC
TGGTGCACTCCTTCCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTTCCTTCCTG
GAGGAGCAGCTGGGGTCCTGCTCCTGCAAGAAAGACTCGTGACTGCCCAGCGCCCCAGGCTG
GACTGAGCCCCTCACGCCGCCCTGCAGCCCCCATGCCCCTGCCCAACATGCTGGGGGTCCAG
AAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCCTCCTCTTCCTCCTCCCC
TTCCTCGGGAGGCTCCCCAGACCCTGGCATGGGATGGGCTGGGATCTTCTCTGTGAATCCAC
CCCTGGCTACCCCCACCCTGGCTACCCCAACGGCATCCCAAGGCCAGGTGGGCCCTCAGCTG
AGGGAAGGTACGAGCTCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCCCGGAG
GCTGGGTGGGGCCTCAGTGGGGGCTGCTGCCTGACCCCCAGCACAATAAAAATGAAACGTGA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGT
CGACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAAT

FIGURE 213

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDTCQSDVDECSARRGGCPQRCINTAGSYWCQCWEGHSLSADGTLCVPKGGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKLQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

FIGURE 214

GCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGCAAG
GGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGCTCC
AGCAGCATCAGAGCAGCCCCTGTGGTTGGCAGCAAAGTTCAGCTTGGCTGGGCCCGCTGTGA
GGGGCTTCGCGCTACGCCCTGCGGTGTCCCGAGGGCTGAGGTCTCCTCATCTTCTCCCTAGC
AGTGGATGAGCAACCCAACGGGGGCCCGGGGAGGGGAACTGGCCCCGAGGGAGAGGAACCCC
AAAGCCACATCTGTAGCCAGGATGAGCAGTGTGAATCCAGGCAGCCCCCAGGACCGGGGAGG
CACAGGTGGCCCCCACCACCCGGAGGAGCAGCTCCTGCCCCTGTCCGGGGGATGACTGATTC
TCCTCCGCCAGGCCACCCAGAGGAGAAGGCCACCCCGCCTGGAGGCACAGGCC**ATG**AGGGGC
TCTCAGGAGGTGCTGCTGATGTGGCTTCTGGTGTTGGCAGTGGGCGGCACAGAGCACGCCTA
CCGGCCCGGCCGTAGGGTGTGTGCTGTCCGGGCTCACGGGGACCCTGTCTCCGAGTCGTTCG
TGCAGCGTGTGTACCAGCCCTTCCTCACCACCTGCGACGGGCACCGGGCCTGCAGCACCTAC
CGAACCATCTATAGGACCGCCTACCGCCGCAGCCCTGGGCTGGCCCCTGCCAGGCCTCGCTA
CGCGTGCTGCCCCGGCTGGAAGAGGACCAGCGGGCTTCCTGGGGCCTGTGGAGCAGCAATAT
GCCAGCCGCCATGCCGGAACGGAGGGAGCTGTGTCCAGCCTGGCCGCTGCCGCTGCCCTGCA
GGATGGCGGGGTGACACTTGCCAGTCAGATGTGGATGAATGCAGTGCTAGGAGGGGCGGCTG
TCCCCAGCGCTGCATCAACACCGCCGGCAGTTACTGGTGCCAGTGTTGGGAGGGGCACAGCC
TGTCTGCAGACGGTACACTCTGTGTGCCCAAGGGAGGGCCCCCCAGGGTGGCCCCCAACCCG
ACAGGAGTGGACAGTGCAATGAAGGAAGAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCT
GGAGGAGAAGCTGCAGCTGGTGCTGGCCCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGC
ATGGGCTCCCGGACCCCGGCAGCCTCCTGGTGCACTCCTTCCAGCAGCTCGGCCGCATCGAC
TCCCTGAGCGAGCAGATTTCCTTCCTGGAGGAGCAGCTGGGGTCCTGCTCCTGCAAGAAAGA
CTCG**TGA**CTGCCCAGCGCTCCAGGCTGGACTGAGCCCCTCACGCCGCCCTGCAGCCCCCATG
CCCCTGCCCAACATGCTGGGGGTCCAGAAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGC
AGGGCCTTCCTCCTCTTCCTCCTCCCCTTCCTCGGGAGGCTCCCCAGACCCTGGCATGGGAT
GGGCTGGGATCTTCTCTGTGAATCCACCCCTGGCTACCCCCACCCTGGCTACCCCAACGGCA
TCCCAAGGCCAGGTGGACCCTCAGCTGAGGGAAGGTACGAGCTCCCTGCTGGAGCCTGGGAC
CCATGGCACAGGCCAGGCAGCCCGGAGGCTGGGTGGGGCCTCAGTGGGGGCTGCTGCCTGAC
CCCCAGCACAATAAAAATGAAACGTG

FIGURE 215

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDTCQSDVDECSARRGGCPQRCINTAGSYWCQCWEGHSLSADGTLCVPKGGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKLQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

FIGURE 216

CCCACGCGTCCGAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGACAGGCCAGGCA
GGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGCAAGGGCTAGGG
TCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGCTCCAGCAGCAT
CAGCAGCCCCCAGGACCGGGGAGGCACAGGTGGCCCCCACCACCCGGAGGAGCAGCTCCTGC
CCCTGTCCGGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGAAGGCCACCCCGC
CTGGAGGCACAGGCC**ATG**AGGGGCTCTCAGGAGGTGCTGCTGATGTGGCTTCTGGTGTTGGC
AGTGGGCGGCACAGAGCACGCCTACCGGCCCGGCCGTAGGGTGTGTGCTGTCCGGGCTCACG
GGGACCCTGTCTCCGAGTCGTTCGTGCAGCGTGTGTACCAGCCCTTCCTCACCACCTGCGAC
GGGCACCGGGCCTGCAGCACCTACCGAACCATCTATAGGACCGCCTACCGCCGCAGCCCTGG
GCTGGCCCCTGCCAGGCCTCGCTACGCGTGCTGCCCCGGCTGGAAGAGGACCAGCGGGCTTC
CTGGGGCCTGTGGAGCAGCAATATGCCAGCCGCCATGCCGGAACGGAGGGAGCTGTGTCCAG
CCTGGCCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAGTCAGATGTGGATGA
ATGCAGTGCTAGGAGGGGCGGCTGTCCCCAGCGCTGCGTCAACACCGCCGGCAGTTACTGGT
GCCAGTGTTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTGTGCCCAAGGGAGGG
CCCCCCAGGGTGGCCCCCAACCCGACAGGAGTGGACAGTGCAATGAAGGAAGAAGTGCAGAG
GCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGCTGGCCCCACTGCACA
GCCTGGCCTCGCAGGCACTGGAGCATGGGCTCCCGGACCCCGGCAGCCTCCTGGTGCACTCC
TTCCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTTCCTTCCTGGAGGAGCAGCT
GGGGTCCTGCTCCTGCAAGAAAGACTCG**TGA**CTGCCCAGCGCCCCAGGCTGGACTGAGCCCC
TCACGCCGCCCTGCAGCCCCCATGCCCCTGCCCAACATGCTGGGGGTCCAGAAGCCACCTCG
GGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCCTCCTCTTCCTCCTCCCCTTCCTCGGGAG
GCTCCCCAGACCCTGGCATGGGATGGGCTGGGATCTTCTCTGTGAATCCACCCCTGGCTACC
CCCACCCTGGCTACCCCAACGGCATCCCAAGGCCAGGTGGGCCCTCAGCTGAGGGAAGGTAC
GAGCTCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCCCGGAGGCTGGGTGGGG
CCTCAGTGGGGGCTGCTGCCTGACCCCCAGCACAATAAAAATGAAACGTG

FIGURE 217

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDTCQSDVDECSARRGGCPQRCVNTAGSYWCQCWEGHSLSADGTLCVPKGGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKLQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

FIGURE 218

GGTTGCCACAGCTGGTTTAGGGCCCCGACCACTGGGGCCCCTTGTCAGGAGGAGACAGCCTC
CCGGCCCGGGGAGGACAAGTCGCTGCCACCTTTGGCTGCCGACGTGATTCCCTGGGACGGTC
CGTTTCCTGCCGTCAGCTGCCGGCCGAGTTGGGTCTCCGTGTTTCAGGCCGGCTCCCCCTTC
CTGGTCTCCCTTCTCCCGCTGGGCCGGTTTATCGGGAGGAGATTGTCTTCCAGGGCTAGCAA
TTGGACTTTTGATGATGTTTGACCCAGCGGCAGGAATAGCAGGCAACGTGATTTCAAAGCTG
GGCTCAGCCTCTGTTTCTTCTCTCGTGTAATCGCAAAACCCATTTTGGAGCAGGAATTCCAA
TCATGTCTGTGATGGTGGTGAGAAAGAAGGTGACACGGAAATGGGAGAAACTCCCAGGCAGG
AACACCTTTTGCTGTGATGGCCGCGTCATGATGGCCCGGCAAAAGGGCATTTTCTACCTGAC
CCTTTTCCTCATCCTGGGGACATGTACACTCTTCTTCGCCTTTGAGTGCCGCTACCTGGCTG
TTCAGCTGTCTCCTGCCATCCCTGTATTTGCTGCCATGCTCTTCCTTTTCTCCATGGCTACA
CTGTTGAGGACCAGCTTCAGTGACCCTGGAGTGATTCCTCGGGCGCTACCAGATGAAGCAGC
TTTCATAGAAATGGAGATAGAAGCTACCAATGGTGCGGTGCCCCAGGGCCAGCGACCACCGC
CTCGTATCAAGAATTTCCAGATAAACAACCAGATTGTGAAACTGAAATACTGTTACACATGC
AAGATCTTCCGGCCTCCCCGGGCCTCCCATTGCAGCATCTGTGACAACTGTGTGGAGCGCTT
CGACCATCACTGCCCCTGGGTGGGGAATTGTGTTGGAAAGAGGAACTACCGCTACTTCTACC
TCTTCATCCTTTCTCTCTCCCTCCTCACAATCTATGTCTTCGCCTTCAACATCGTCTATGTG
GCCCTCAAATCTTTGAAAATTGGCTTCTTGGAGACATTGAAAGAAACTCCTGGAACTGTTCT
AGAAGTCCTCATTTGCTTCTTTACACTCTGGTCCGTCGTGGGACTGACTGGATTTCATACTT
TCCTCGTGGCTCTCAACCAGACAACCAATGAAGACATCAAAGGATCATGGACAGGGAAGAAT
CGCGTCCAGAATCCCTACAGCCATGGCAATATTGTGAAGAACTGCTGTGAAGTGCTGTGTGG
CCCCTTGCCCCCCAGTGTGCTGGATCGAAGGGGTATTTTGCCACTGGAGGAAAGTGGAAGTC
GACCTCCCAGTACTCAAGAGACCAGTAGCAGCCTCTTGCCACAGAGCCCAGCCCCCACAGAA
CACCTGAACTCAAATGAGATGCCGGAGGACAGCAGCACTCCCGAAGAGATGCCACCTCCAGA
GCCCCCAGAGCCACCACAGGAGGCAGCTGAAGCTGAGAAGTAGCCTATCTATGGAAGAGACT
TTTGTTTGTGTTTAATTAGGGCTATGAGAGATTTCAGGTGAGAAGTTAAACCTGAGACAGAG
AGCAAGTAAGCTGTCCCTTTTAACTGTTTTTCTTTGGTCTTTAGTCACCCAGTTGCACACTG
GCATTTTCTTGCTGCAAGCTTTTTTAAATTTCTGAACTCAAGGCAGTGGCAGAAGATGTCAG
TCACCTCTGATAACTGGAAAAATGGGTCTCTTGGGCCCTGGCACTGGTTCTCCATGGCCTCA
GCCACAGGGTCCCCTTGGACCCCCTCTCTTCCCTCCAGATCCCAGCCCTCCTGCTTGGGGTC
ACTGGTCTCATTCTGGGGCTAAAAGTTTTTGAGACTGGCTCAAATCCTCCCAAGCTGCTGCA
CGTGCTGAGTCCAGAGGCAGTCACAGAGACCTCTGGCCAGGGGATCCTAACTGGGTTCTTGG
GGTCTTCAGGACTGAAGAGGAGGGAGAGTGGGGTCAGAAGATTCTCCTGGCCACCAAGTGCC
AGCATTGCCCACAAATCCTTTTAGGAATGGGACAGGTACCTTCCACTTGTTGTANNNNNNNN
NNNNNNNNNNNNNNNNNNTTGTTTTTCCTTTTGACTCCTGCTCCCATTAGGAGCAGGAATG
GCAGTAATAAAAGTCTGCACTTTGGTCATTTCTTTTCCTCAGAGGAAGCCCGAGTGCTCACT
TAAACACTATCCCCTCAGACTCCCTGTGTGAGGCCTGCAGAGGCCCTGAATGCACAAATGGG
AAACCAAGGCACAGAGAGGCTCTCCTCTCCTCTCCTCTCCCCCGATGTACCCTCAAAAAAAA
AAAAATGCTAACCAGTTCTTCCATTAAGCCTCGGCTGAGTGAGGGAAAGCCCAGCACTGCTG
CCCTCTCGGGTAACTCACCCTAAGGCCTCGGCCCACCTCTGGCTATGGTAACCACACTGGGG
GCTTCCTCCAAGCCCCGCTCTTCCAGCACTTCCACCGGCAGAGTCCCAGAGCCACTTCACCC
TGGGGGTGGGCTGTGGCCCCCAGTCAGCTCTGCTCAGGACCTGCTCTATTTCAGGGAAGAAG
ATTTATGTATTATATGTGGCTATATTTCCTAGAGCACCTGTGTTTTCCTCTTTCTAAGCCAG
GGTCCTGTCTGGATGACTTATGCGGTGGGGGAGTGTAAACCGGAACTTTTCATCTATTTGAA
GGCGATTAAACTGTGTCTAATGCA

FIGURE 219

MSVMVVRKKVTRKWEKLPGRNTFCCDGRVMMARQKGIFYLTLFLILGTCTLFFAFECRYLAV
QLSPAIPVFAAMLFLFSMATLLRTSFSDPGVIPRALPDEAAFIEMEIEATNGAVPQGQRPPP
RIKNFQINNQIVKLKYCYTCKIFRPPRASHCSICDNCVERFDHHCPWVGNCVGKRNYRYFYL
FILSLSLLTIYVFAFNIVYVALKSLKIGFLETLKETPGTVLEVLICFFTLWSVVGLTGFHTF
LVALNQTTNEDIKGSWTGKNRVQNPYSHGNIVKNCCEVLCGPLPPSVLDRRGILPLEESGSR
PPSTQETSSSLLPQSPAPTEHLNSNEMPEDSSTPEEMPPPEPPEPPQEAAEAEK

FIGURE 220

AAAACCCTGTATTTTTTACAATGCAAATAGACAATNANCCTGGAGGTCTTTGAATTAGGTAT
TATAGGGATGGTGGGGTTGATTTTTNTTCCTGGAGGCTTTTGGCTTTGGACTCTCNCTTTCT
CCCACAGAGCNCTTCGACCATCACTGCCCCTGGGTGGGGAATTGTGTTGGAAAGAGGAACTA
CCGCTANTTCTACCTCTTCATCCTTTNTCTCTCCCNCCTCACAATCTATGTCTTCGCCTTCA
ACATCGT

FIGURE 221

GTTGTGTCCTTCAGCAAAACAGTGGATTTAAATCTCCTTGCACAAGCTTGAGAGCAACACAA
TCTATCAGGAAAGAAAGAAAGAAAAAAACCGAACCTGACAAAAAAGAAGAAAAAGAAGAAGA
AAAAAAATC**ATG**AAAACCATCCAGCCAAAAATGCACAATTCTATCTCTTGGGCAATCTTCAC
GGGGCTGGCTGCTCTGTGTCTCTTCCAAGGAGTGCCCGTGCGCAGCGGAGATGCCACCTTCC
CCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGTGCACTATT
GACAACCGGGTCACCCGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTATGCTGGGAATGA
CAAGTGGTGCCTGGATCCTCGCGTGGTCCTTCTGAGCAACACCCAAACGCAGTACAGCATCG
AGATCCAGAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTCGGTGCAGACAGACAAC
CACCCAAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCCAAAATTGTAGAGATTTC
TTCAGATATCTCCATTAATGAAGGGAACAATATTAGCCTCACCTGCATAGCAACTGGTAGAC
CAGAGCCTACGGTTACTTGGAGACACATCTCTCCCAAAGCGGTTGGCTTTGTGAGTGAAGAC
GAATACTTGGAAATTCAGGGCATCACCCGGGAGCAGTCAGGGGACTACGAGTGCAGTGCCTC
CAATGACGTGGCCGCGCCCGTGGTACGGAGAGTAAAGGTCACCGTGAACTATCCACCATACA
TTTCAGAAGCCAAGGGTACAGGTGTCCCCGTGGGACAAAAGGGGACACTGCAGTGTGAAGCC
TCAGCAGTCCCCTCAGCAGAATTCCAGTGGTACAAGGATGACAAAAGACTGATTGAAGGAAA
GAAAGGGGTGAAAGTGGAAAACAGACCTTTCCTCTCAAAACTCATCTTCTTCAATGTCTCTG
AACATGACTATGGGAACTACACTTGCGTGGCCTCCAACAAGCTGGGCCACACCAATGCCAGC
ATCATGCTATTTGGTCCAGGCGCCGTCAGCGAGGTGAGCAACGGCACGTCGAGGAGGGCAGG
CTGCGTCTGGCTGCTGCCTCTTCTGGTCTTGCACCTGCTTCTCAAATTT**TGA**TGTGAGTGCC
ACTTCCCCACCCGGGAAAGGCTGCCGCCACCACCACCACCAACACAACAGCAATGGCAACAC
CGACAGCAACCAATCAGATATATACAAATGAAATTAGAAGAAACACAGCCTCATGGGACAGA
AATTTGAGGGAGGGGAACAAAGAATACTTTGGGGGGAAAAGAGTTTTAAAAAAGAAATTGAA
AATTGCCTTGCAGATATTTAGGTACAATGGAGTTTTCTTTTCCCAAACGGGAAGAACACAGC
ACACCCGGCTTGGACCCACTGCAAGCTGCATCGTGCAACCTCTTTGGTGCCAGTGTGGGCAA
GGGCTCAGCCTCTCTGCCCACAGAGTGCCCCCACGTGGAACATTCTGGAGCTGGCCATCCCA
AATTCAATCAGTCCATAGAGACGAACAGAATGAGACCTTCCGGCCCAAGCGTGGCGCTGCGG
GCACTTTGGTAGACTGTGCCACCACGGCGTGTGTTGTGAAACGTGAAATAAAAAGAGCAAAA
AAAAA

FIGURE 222

MKTIQPKMHNSISWAIFTGLAALCLFQGVPVRSGDATFPKAMDNVTVRQGESATLRCTIDNR
VTRVAWLNRSTILYAGNDKWCLDPRVVLLSNTQTQYSIEIQNVDVYDEGPYTCSVQTDNHPK
TSRVHLIVQVSPKIVEISSDISINEGNNISLTCIATGRPEPTVTWRHISPKAVGFVSEDEYL
EIQGITREQSGDYECSASNDVAAPVVRRVKVTVNYPPYISEAKGTGVPVGQKGTLQCEASAV
PSAEFQWYKDDKRLIEGKKGVKVENRPFLSKLIFFNVSEHDYGNYTCVASNKLGHTNASIML
FGPGAVSEVSNGTSRRAGCVWLLPLLVLHLLLKF

FIGURE 223

GAAAAAAAATCATGAAAACCATCCAGCCAAAAATGCACAATTCTATCTCTTGGGCAATCTTC
ACGGGGCTGGCTGCTCTGTGTCTCTTCCAAGGAGTGCCCGTGCGCAGCGGAGATGCCACCTT
CCCCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGTGCACTA
TTGACAACCGGGTCACCCGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTATGCTGGGAAT
GACAAGTGGTGCCTGGATCCTCGCGTGGTCCTTCTGAGCAACACCCAAACGCAGTACAGCAT
CGAGATCCAGAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTCGGTGCAGACAGACA
ACCACCCAAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCCAAAATTGTAGAGATT
TCTTCAGATATCTCCATTAATGAAGGGAACAATATTAGCCTCACCTGCATAGCAACTGGTAG
ACCAGAG

FIGURE 224

ATGGCTGGTGACGGCGGGGCCGGGCAGGGGACCGGGGCCGCGGCCCGGGAGCGGGCCAGCTG
CCGGGAGCCCTGAATCACCGCCTGGCCCGACTCCACC**ATG**AACGTCGCGCTGCAGGAGCTGG
GAGCTGGCAGCAACGTGGGATTCCAGAAGGGGACAAGACAGCTGTTAGGCTCACGCACGCAG
CTGGAGCTGGTCTTAGCAGGTGCCTCTCTACTGCTGGCTGCACTGCTTCTGGGCTGCCTTGT
GGCCCTAGGGGTCCAGTACCACAGAGACCCATCCCACAGCACCTGCCTTACAGAGGCCTGCA
TTCGAGTGGCTGGAAAAATCCTGGAGTCCCTGGACCGAGGGGTGAGCCCCTGTGAGGACTTT
TACCAGTTCTCCTGTGGGGGCTGGATTCGGAGGAACCCCCTGCCCGATGGGCGTTCTCGCTG
GAACACCTTCAACAGCCTCTGGGACCAAAACCAGGCCATACTGAAGCACCTGCTTGAAAACA
CCACCTTCAACTCCAGCAGTGAAGCTGAGCAGAAGACACAGCGCTTCTACCTATCTTGCCTA
CAGGTGGAGCGCATTGAGGAGCTGGGAGCCCAGCCACTGAGAGACCTCATTGAGAAGATTGG
TGGTTGGAACATTACGGGGCCCTGGGACCAGGACAACTTTATGGAGGTGTTGAAGGCAGTAG
CAGGGACCTACAGGGCCACCCCATTCTTCACCGTCTACATCAGTGCCGACTCTAAGAGTTCC
AACAGCAATGTTATCCAGGTGGACCAGTCTGGGCTCTTTCTGCCCTCTCGGGATTACTACTT
AAACAGAACTGCCAATGAGAAAGTGCTCACTGCCTATCTGGATTACATGGAGGAACTGGGGA
TGCTGCTGGGTGGGCGGCCCACCTCCACGAGGGAGCAGATGCAGCAGGTGCTGGAGTTGGAG
ATACAGCTGGCCAACATCACAGTGCCCCAGGACCAGCGGCGCGACGAGGAGAAGATCTACCA
CAAGATGAGCATTTCGGAGCTGCAGGCTCTGGCGCCCTCCATGGACTGGCTTGAGTTCCTGT
CTTTCTTGCTGTCACCATTGGAGTTGAGTGACTCTGAGCCTGTGGTGGTGTATGGGATGGAT
TATTTGCAGCAGGTGTCAGAGCTCATCAACCGCACGGAACCAAGCATCCTGAACAATTACCT
GATCTGGAACCTGGTGCAAAAGACAACCTCAAGCCTGGACCGACGCTTTGAGTCTGCACAAG
AGAAGCTGCTGGAGACCCTCTATGGCACTAAGAAGTCCTGTGTGCCGAGGTGGCAGACCTGC
ATCTCCAACACGGATGACGCCCTTGGCTTTGCTTTGGGGTCACTCTTCGTGAAGGCCACGTT
TGACCGGCAAAGCAAAGAAATTGCAGAGGGGATGATCAGCGAAATCCGGACCGCATTTGAGG
AGGCCCTGGGACAGCTGGTTTGGATGGATGAGAAGACCCGCCAGGCAGCCAAGGAGAAAGCA
GATGCCATCTATGATATGATTGGTTTCCCAGACTTTATCCTGGAGCCCAAAGAGCTGGATGA
TGTTTATGACGGGTACGAAATTTCTGAAGATTCTTTCTTCCAAAACATGTTGAATTTGTACA
ACTTCTCTGCCAAGGTTATGGCTGACCAGCTCCGCAAGCCTCCCAGCCGAGACCAGTGGAGC
ATGACCCCCCCAGACAGTGAATGCCTACTACCTTCCAACTAAGAATGAGATCGTCTTCCCCGC
TGGCATCCTGCAGGCCCCCTTCTATGCCCGCAACCACCCCAAGGCCCTGAACTTCGGTGGCA
TCGGTGTGGTCATGGGCCATGAGTTGACGCATGCCTTTGATGACCAAGGGCGCGAGTATGAC
AAAGAAGGGAACCTGCGGCCCTGGTGGCAGAATGAGTCCCTGGCAGCCTTCCGGAACCACAC
GGCCTGCATGGAGGAACAGTACAATCAATACCAGGTCAATGGGGAGAGGCTCAACGGCCGCC
AGACGCTGGGGGGAGAACATTACTGACAACGGGGGGCTGAAGGCTGCCTACAATGCTTACAAA
GCATGGCTGAGAAAGCATGGGGAGGAGCAGCAACTGCCAGCCGTGGGGCTCACCAACCACCA
GCTCTTCTTCGTGGGATTTGCCCAGGTGTGGTGCTCGGTCCGCACACCAGAGAGCTCTCACG
AGGGGCTGGTGACCGACCCCCACAGCCCTGCCCGCTTCCGCGTGCTGGGCACTCTCTCCAAC
TCCCGTGACTTCCTGCGGCACTTCGGCTGCCCTGTCGGCTCCCCCATGAACCCAGGGCAGCT
GTGTGAGGTGTGG**TAG**ACCTGGATCAGGGGAGAAATGGCCAGCTGTCACCAGACCTGGGGCA
GCTCTCCTGACAAAGCTGTTTGCTCTTGGGTTGGGAGGAAGCAAATGCAAGCTGGGCTGGGT
CTAGTCCCTCCCCCCCACAGGTGACATGAGTACAGACCCTCCTCAATCACCACATTGTGCCT
CTGCTTTGGGGGTGCCCCTGCCTCCAGCAGAGCCCCCACCATTCACTGTGACATCTTTCCGT
GTCACCCTGCCTGGAAGAGGTCTGGGTGGGGAGGCCAGTTCCCATAGGAAGGAGTCTGCC

FIGURE 225

MNVALQELGAGSNVGFQKGTRQLLGSRTQLELVLAGASLLLAALLLGCLVALGVQYHRDPSH
STCLTEACIRVAGKILESLDRGVSPCEDFYQFSCGGWIRRNPLPDGRSRWNTFNSLWDQNQA
ILKHLLENTTFNSSSEAEQKTQRFYLSCLQVERIEELGAQPLRDLIEKIGGWNITGPWDQDN
FMEVLKAVAGTYRATPFFTVYISADSKSSNSNVIQVDQSGLFLPSRDYYLNRTANEKVLTAY
LDYMEELGMLLGGRPTSTREQMQQVLELEIQLANITVPQDQRRDEEKIYHKMSISELQALAP
SMDWLEFLSFLLSPLELSDSEPVVVYGMDYLQQVSELINRTEPSILNNYLIWNLVQKTTSSL
DRRFESAQEKLLETLYGTKKSCVPRWQTCISNTDDALGFALGSLFVKATFDRQSKEIAEGMI
SEIRTAFEEALGQLVWMDEKTRQAAKEKADAIYDMIGFPDFILEPKELDDVYDGYEISEDSF
FQNMLNLYNFSAKVMADQLRKPPSRDQWSMTPQTVNAYYLPTKNEIVFPAGILQAPFYARNH
PKALNFGGIGVVMGHELTHAFDDQGREYDKEGNLRPWWQNESLAAFRNHTACMEEQYNQYQV
NGERLNGRQTLGENITDNGGLKAAYNAYKAWLRKHGEEQQLPAVGLTNHQLFFVGFAQVWCS
VRTPESSHEGLVTDPHSPARFRVLGTLSNSRDFLRHFGCPVGSPMNPGQLCEVW

FIGURE 226A

GCCCGGCCCTCCGCCCTCCGCACTCCCGCCTCCCTCCCTCCGCCCGCTCCCGCGCCCTCCTC
CCTCCCTCCTCCCCAGCTGTCCCGTTCGCGTCATGCCGAGCCTCCCGGCCCCGCCGGCCCCG
CTGCTGCTCCTCGGGCTGCTGCTGCTCGGCTCCCGGCCGGCCCGCGGCGCCGGCCCAGAGCC
CCCCGTGCTGCCCATCCGTTCTGAGAAGGAGCCGCTGCCCGTTCGGGGAGCGGCAGGTAGGT
GGGCGCCCGGGGGAGGCGCGGGCGGGGAGTCGGGCTCGGGGCGAGTCAGCGCCAGCCCGGAG
GGGGCGCGGGGCGCAGGTGGCTCGGCGCGGCGGGCGGCCCGGAGGGTGGGCGGGGGCAGAAG
GGCGCGGTGCCTGGGACCCGGGACCCGCGGGCAGCCCCCGGGGCGGCACACGGCGCGAGCTG
GGCAGCGGCCTCCAGCCAAGCCCGTCCCCGCAGGCTGCACCTTCGGCGGGAAGGTCTATGCC
TTGGACGAGACGTGGCACCCGGACCTAGGGGAGCCATTCGGGGTGATGCGCTGCGTGCTGTG
CGCCTGCGAGGCGCAGTGGGGTCGCCGTACCAGGGGCCCTGGCAGGGTCAGCTGCAAGAACA
TCAAACCAGAGTGCCCAACCCCGGCCTGTGGGCAGCCGCGCCAGCTGCCGGGACACTGCTGC
CAGACCTGCCCCCAGGACTTCGTGGCGCTGCTGACAGGGCCGAGGTCGCAGGCGGTGGCACG
AGCCCGAGTCTCGCTGCTGCGCTCTAGCCTCCGCTTCTCTATCTCCTACAGGCGGCTGGACC
GCCCTACCAGGATCCGCTTCTCAGACTCCAATGGCAGTGTCCTGTTTGAGCACCCTGCAGCC
CCCACCCAAGATGGCCTGGTCTGTGGGGTGTGGCGGGCAGTGCCTCGGTTGTCTCTGCGGCT
CCTTAGGGCAGAACAGCTGCATGTGGCACTTGTGACACTCACTCACCCTTCAGGGGAGGTCT
GGGGGCCTCTCATCCGGCACCGGGCCCTGTCCCCAGAGACCTTCAGTGCCATCCTGACTCTA
GAAGGCCCCCACCAGCAGGGCGTAGGGGGCATCACCCTGCTCACTCTCAGTGACACAGAGGA
CTCCTTGCATTTTTTGCTGCTCTTCCGAGGCCTTGCAGGACTAACCCAGGTTCCCTTGAGGC
TCCAGATTCTACACCAGGGGCAGCTACTGCGAGAACTTCAGGCCAATGTCTCAGCCCAGGAA
CCAGGCTTTGCTGAGGTGCTGCCCAACCTGACAGTCCAGGAGATGGACTGGCTGGTGCTGGG
GGAGCTGCAGATGGCCCTGGAGTGGGCAGGCAGGCCAGGGCTGCGCATCAGTGGACACATTG
CTGCCAGGAAGAGCTGCGACGTCCTGCAAAGTGTCCTTTGTGGGGCTAATGCCCTGATCCCA
GTCCAAACGGGTGCTGCCGGCTCAGCCAGCCTCACTCTGCTAGGAAATGGCNCCCTGATCCT
CCAGGTGCAATTGGTAGGGACAACCAGTGAGGTGGTGGCCATGACACTGGAAACCAAGCCTC
AGCGGAGGGATCAGCCCACTGTCCTGTGCCACATGGCTGGCCTATCCTCCCCTGCCCCCAGG
CCGTGGGTATCTGCCCTGGGCTGGGGTGCCCGAGGGGCTCATATGCTGCTGCAGAATGAGCT
CTTCCTGAACGTGGGCACCAAGGACTTCCCAGACGGAGAGCTTCGGGGGCAACGTGGCTGCC
CTGCCCTACTGTGGGGCATAGCGCCCGCCCTGCCCGTGCCCCTAGCAGGAGCCCTGGTGCTA
CCCCCTGTGAAGAGCCAAGCAGCAGGGCACGCCTGGCTTTCCTTGGATACCCACTGTCACCT
GCACTATGAAGTGCTGCTGGCTGGGCTTGGTGGCTCAGAACAAGGCACTGTCACTGCCCACC
TCCTTGGGCCTCCTGGAACGCCAGGGCCTCGGCGGCTGCTGAAGGGATTCTATGGCTCAGAG
GCCCAGGGTGTGGTGAAGGACCTGGAGCCGGAACTGCTGCGGCACCTGGCAAAAGGCATGGC
TTCCCTGATGATCACCACCAAGGTAGCCCCAGAGGGGAGCTCCGAGGGCAGCCTCTCCTCCC
AGGTGCACATAGCCAACCAATGTGAGGTTGGCGGACTGCGCCTGGAGGCGGCCGGGGCCGAG
GGGGTGCGGGCGCTGGGGGCTCCGGATACAGCCTCTGCTGCGCCGCCTGTGGTGCCTGGTCT
CCCGGCCCTAGCGCCCGCCAAACCTGGTGGTCCTGGGCGGCCCCGAGACCCCAACACATGCT
TCTTCGAGGGGCAGCAGCGCCCCCACGGGGCTCGCTGGGCGCCCAACTACGACCCGCTCTGC
TCACTCTGCACCTGCCAGAGACGAACGGTGATCTGTGACCCGGTGGTGTGCCCACCGCCCAG
CTGCCCACACCCGGTGCAGGCTCCCGACCAGTGCTGCCCTGTTTGCCCTGGCTGCTATTTTG
ATGGTGACCGGAGCTGGCGGGCAGCGGGTACGCGGTGGCACCCCGTTGTGCCCCCCTTTGGC
TTAATTAAGTGTGCTGTCTGCACCTGCAAGCAGGGGGGCACTGGAGAGGTGCACTGTGAGAA
GGTGCAGTGTCCCCGGCTGGCCTGTGCCCAGCCTGTGCGTGTCAACCCCACCGACTGCTGCA
AACAGTGTCCAGGTGAGGCCCACCCCCAGCTGGGGGACCCCATGCAGGCTGATGGGCCCCGG
GGCTGCCGTTTTGCTGGGCAGTGGTTCCCAGAGAGTCAGAGCTGGCACCCCTCAGTGCCCCC
GTTTGGAGAGATGAGCTGTATCACCTGCAGATGTGGGGTAAGTGGGGAGCAGAGGCTTGTGT
GAGGTGGGTACTGGGAGCCTGGTCTGGAGTAGGGAGACCTTCCCAGGGAGGTCCCTGAAGAA
GCTGAAGGTCACTGTGTCCCAGTGCCTCTGGGGGACACTCAGTGTCTGCTCTGTCTTGTACC
AGGCAGGGGTGCCTCACTGTGAGCGGGATGACTGTTCACTGCCACTGTCCTGTGGCTCGGGG
AAGGAGAGTCGATGCTGTTCCCGCTGCACGGCCCACCGGCGGCGTAAGTGAGGGAGTCCAGG
GTCAGCAGCTGTGAGTGGAGGGCTCACCTGCCTGTGGGACTCCTGATCAGGGAAGGGAGCAC
TCACTGTGTGCAGGAACAGTGCAGCCTGCCTCACAAGTGCCATTCCAATCCACCCTCACAGC
AACCTGGTGGAATTGTTATTTATGACCTTTTCTTTACAAATGAGATTTCTGAAGCTCAGAGA
AATTAAGCAACGAGATGAAGGTCACCCAGCTGTGTGCACTGACCTGTTTAGAAAATACTGGC

FIGURE 226B

CTTTCTGGGACCAAGGCAGGGATGCTTTGCCCTGCCCTCTATGCCTCTCTGTGCCTCTCCAC
TCCCTCTCCCCTCCTCCAACATTCCCTCCCTTCTGTCTCCAGCAGCCCCAGAGACCAGAACT
GATCCAGAGCTGGAGAAAGAAGCCGAAGGCTCTTAGGGAGCAGCCAGAGGGCCAAGTGACCA
AGAGGATGGGGCCTGAGCTGGGGAAGGGGTGGCATCGAGGACCTTCTTGCATTCTCCTGTGG
GAAGCCCAGTGCCTTTGCTCCTCTGTCCTGCCTCTACTCCCACCCCCACTACCTCTGGGAAC
CACAGCTCCACAAGGGGGAGAGGCAGCTGGGCCAGACCGAGGTCACAGCCACTCCAAGTCCT
GCCCTGCCACCCTCGGCCTCTGTCCTGGAAGCCCCACCCCTTTCTTCCTGTACATAATGTCA
CTGGCTTGTTGGGATTTTTAATTTATCTTCACTCAGCACCAAGGGCCCCGGACACTCCACTC
CTGCTGCCCCTGAGCTGAGCAGAGTCATTATTGGAGAGTTTTGTATTTATTAAAACATTTCT
TTTTCAGTCTTTGGGCATGAGGTTGGCTCTTTGTGGCCAGGAACCTGAGTGGGGCCTGGTGG
AGAAGGGGCNGAGAGTAGGAGGTGAGAGAGAGGAGCTCTGACACTTGGGGAGCTGAAAGAGA
CCTGGAGAGGCAGAGGATAGCGTGGCNNTTGGCTGGCATNCCTGGGTTCCGCAGAGGGGCTG
GGGATGGTTCTTGAGATGGTCTAGAGACTCAAGAATTTAGGGAAGTAGAAGCAGGATTTTGA
CTCAAGTTTAGTTTCCCACATCGCTGGCCTGTTTGCTGACTTCATGTTTGAAGTTGCTCCAG
AGAGAGAATCAAAGGTGTCACCAGCCCCTCTCTCCCTCCTTCCCTTCCCTTCCCTTTCTTTC
CCTCCCCTCCCCTCCCCTCCCCTCCCCTCC

FIGURE 227

GGCCGAGCGGGGGTGCTGCGCGGCGGCCGTGATGGCTGGTGACGGCGGGGCCGGGCAGGGGA
CCGGGGCCGCGGCCCGGGAGCGGGCCAGCTGCCGGGAGCCCTGAATCACCGCCTGGCCCGAC
TCCACCATGAACGTCGCGCTGCAGGAGCTGGGAGCTGGCAGCAACGTGGGATTCCAGAAGGG
GACAAGACAGCTGTTAGGCTCACGCACGCAGCTGGAGCTGGTCTTAGCAGGTGCCTCTCTAC
TGCTGGCTGCACTGCTTCTGGGCTGCCTTGTGGCCCTAGGGGTCCAGTACCACAGAGACCCA
TCCCACAGCACCTGCCTTACAGAGGCCTGCATTCGAGTGGCTGGAAAAATCCTGGAGTCCCT
GGACCGAGGGGTGAGCCCCTGTGAGGACTTTTACCAGTTCTCCTGTGGGGGCTGGATTCGGA
GGAACCCCCTGCCCGATGGGCGTTCTCGCTGGAACACCTTCAACAGCCTCTGGGACCAAAAC
CAGGCCATACTGAAGCACCTGCTTGAAAACACCACCTTCAACTCCAGCAGTGAAGCTGAGCA
GAAGACACAGCGCTTCTACCTATCTTGCCTACAGGTGGAGCGCATTGAGGAGCTGGGAGCCC
AGCCACTGAGAGACCTCATTGAGAAGATTGGTGGTTGGAACATTACGGGGCCCTGGGACCAG
GACAACTTTATGGAGGTGTTGAAGGCAGTAGCAGGGACCTACAGGGCCACCCCATTCTTCAC
CGTCTACATCAGTGCCGACTCTAAGAGTTCCAACAGCAATGTTATCCAGGTGGACCAGTCTG
GGCTCTTTCTGCCCTCTCGGGATTACTACTTAAACAGAACTGCCAATGAGAAAGTAAGGAAC
ATCTTCCGAACCCCCATCCCTACCCCTGGCTGAGCTGGGCTGATCCCTGTTGACTTTTCCCT
TTGCCAAGGGTCAGAGCAGGGAAGGTGAGCCTATCCTGTCACCTAGTGAACAAACTGCCCCT
CCTTTCTTTCTTCTTTTCTTCCTCCCTCCCTCCCTTTCTTCCCCTTTTCCTTCCTTCCTTCC
TCTTATTCTTCTAGTAGGTTTCATAGACACCTACTGTGTGCCAGGTCCAGTGGGGGAATTCG
GAGATATAAGTTTCCGAGCCATTGCCACAGGAAGCGTTCAGTGTCGATGGGTTCATGGACCT
AGATAGGCTGATAACAAAGCTCACAAGAGGGTCCTGAGGATTCAGGAGAGACTTATGGAGCC
AGCAAAGTCTTCCTGAAGAGATTGCATTTGAGCCAGGTCCTGTAG

FIGURE 228

ATGCCTACTACCTTCCAACTAAGAATGAGATCGTCTTCCCCGCTGGCATCCTGCAGGCCCCC
TTCTATGCCCGCAACCACCCCAAGGCCCTGAACTTCGGTGGCATCGGTGTGGTCATGGGCCA
TGAGTTGACGCATGCCTTTGATGACCAAGGGCGCGAGTATGACAAAGAAGGGAACCTGCGGC
CCTGGTGGCAGAATGAGTCCCTGGCAGCCTTCCGGAACCACACGGCCTGCATGGAGGAACAG
TACAATCAATACCAGGTCAATGGGGAGAGGCTCAACGGCCGCCAGACGCTGGGGGAGAACAT
TGCTGACAACGGGGGGCTGAAGGCTGCCTACAATGCTTACAAAGCATGGCTGAGAAAGCATG
GGGAGGAGCAGCAACTGCCAGCCGTGGGGCTCACCAACCACCAGCTCTTCTTCGTGGGATTT
GCCCAGGTGTGGTGCTCGGTCCGCACACCAGAGAGCTCTCACGAGGGGCTGGTGACCGACCC
CCACAGCCCTGCCCGCTTCCGCGTGCTGGGCACTCTCTCCAACTCCCGTGACTTCCTGCGGC
ACTTCGGCTGCCCTGTCGGCTCCCCCATGAACCCAGGGCAGCTGTGTGAGGTGTGGTAGACC
TGGATCAGGGGAGAAATGGCCAGCTGTCACCAGACCTGGGGCAGCTCTCCTGACAAAGCTGT
TTGCTCTTGGGTTGGGAGGAAGCAAATGCAAGCTGGGCTGGGTCTAGTCCCTCCCCCCCACA
GGTGACATGAGTACAGACCCTCCTCAATCACCACATTGTGCCTCTGCTTTGGGGGTGCCCCT
GCCTCCAGCAGAGCCCCCACCATTCACTGTGACATCTTTCCGTGTCACCCTGCCTGGAAGAG
GTCTGGGTGGGGAGGCCAGTTCCCATAGGAAGGAGTCTGCCTCTTCTGTCCCCAGGCTCACT
CAGCCTGGCGGCCATGGGGCCTGCCGTGCCTGCCCCACTGTGACCCACAGGCCTGGGTGGTG
TACCTCCTGGACTTCTCCCCAGGCTCACTCAGTGCGCACTTAGGGGTGGACTCAGCTCTGTC
TGGCTCACCCTCACGGGCTACCCCCACCTCACCCTGTGCTCCTTGTGCCACTGCTCCCAGTG
CTGCTGCTGACCTTCACTGACAGCTCCTAGTGGAAGCCCAAGGGCCTCTGAAAGCCTCCTGC
TGCCCACTGTTTCCCTGGGCTGAGAGGGGAAGTGCATATGTGTAGCGGGTACTGGTTCCTGT
GTCTTAGGGCACAAGCCTTAGCAAATGATTGATTCTCCCTGGACAAAGCAGGAAAGCAGATA
GAGCAGGGAAAAGGAAGAACAGAGTTTATTTTTACAGAAAAGAGGGTGGGAGGGTGTGGTCT
TGGCCCTTATAGGACC